序

这本书的作者——欧阳维诚、张垚、肖果能，是我十分熟悉的三位数学教授，他们各自都有非常优秀的工作，有极高的数学修养。

说到数学修养，不只是指对数学知识的掌握，更包括对数学精神的体验，对数学文化的感受，以及他们对数学教育事业的执著。

数学并不是由一串串的计算公式和一系列定理堆砌而成的，数学的精华在于它的思想、精神和理想，正是它的这种精华给了人类文化演进以极其巨大的影响。三位教授著作的选材正在于表现这种精华。

比较丰富的数学知识是形成良好数学修养的基础，而数学修养的增强将使我们能更好地理解数学，从而又促进我们增长和充实数学知识。所以阅读一些关于数学思想、数学方法、数学史的著作是大有益处的。

三位教授撰写的《初等数学思想方法选讲》一书还有其独到之处。

首先，虽然关于数学思想方法方面的著述已陆续出版了一些，但像这部著作这样将数学内容、数学思想、数学方法、数学本质、数学精神以及数学历史有机地融汇在一起，其论说又这样简炼，实属罕见。

其次，虽然三位教授涉及的知识面很宽，但基本上是以初等数学为背景材料，即便涉及一些高等内容也尽量采用初等方

法或比较通俗的叙述方式，因而其可读性会更强，相信读者面也会比较宽。

再次，三位教授一直非常关注数学教育，又有丰富的实践经验，因而他们的叙说更贴近实际，数学爱好者、数学专业的学生、某些非数学专业的学生以及数学教师都有可能通过这本书而有新的体验和感受。

相信三位教授是又做了一件大有益于数学研究和数学教育的工作。

张楚廷

2000 年元月

前　言

秘鲁大学教授 J.N.Kapur 曾在他的《数学家谈数学本质》一书的序言中提出这样一个观点：数学本质是每一位数学教员，每一位学习数学的学生，每一位受过教育的人，每一位科学(包括社会科学)技术工作者，甚至每一位学生家长都关心的问题，是世界上几乎每个人都应该感兴趣的问题。

我们认为，把他所说的"数学本质"换成"数学思想方法"也应该如此。

每一位数学教员都应该对数学思想方法感兴趣，因为如果他只理解数学知识而不理解数学的思想方法，那是很难教好这门课程的。

每一位数学专业的学生肯定都对数学思想方法感兴趣。一个人的精力有限，数学的知识无穷，只有掌握了数学思想方法，才能深刻地理解数学、欣赏数学和应用数学。

每一个受过教育的人也应该对数学思想方法感兴趣，因为他多少学过一些数学，无论他是否喜爱数学，是否应用数学，但数学思想方法对他从事的专业一定有所帮助。

每一位科学工作者都应该对数学思想方法感兴趣，因为数学已经向各种科学(包括社会科学)渗透，在他的研究工作中肯定会遇到数学思想方法的挑战。

湖南教育出版社郑绍辉先生多年从事数学教育图书的编辑出版工作，卓有成效。他有鉴于对数学教师、青年学生提供一本数学思想方法方面的读物的重要性，几年前就提出了这个图

书选题。

我们愉快地接受了原来自以为能够勉为其难的任务，但是"事非经过不知难"，具体动起笔来，却碰到了许多意想不到的困难。

困难主要来自两个方面：

首先，什么是数学思想，什么又是数学方法，还没有大家一致认同的明确的界定，名不正则言不顺，这是第一个困难。

其次，数学思想方法内容庞杂，纵横交错，要提纲挈领而理不清头绪，想条陈缕析而把不住脉络，这是第二个困难。

困难当前的情况下，我们只好采取了我行我素的对策，按照我们自己对数学思想方法的粗浅理解去展开我们的讨论。

首先，我们把数学思想分为三类或三个层次：第一是对数学本质的认识方面的思想，它回答"数学是什么"的问题。第二是数学的地位、作用、研究和发展方面的思想，它回答"数学向何处去"的问题。第三是解决具体数学问题的思想，它回答"数学怎样论证"的问题。根据责任编辑的组稿意图，本书主要是讨论第三层次的数学思想，为了方便，我们把它简称为数学解题思想。

其次，我们认为，数学解题思想虽然内容广泛、错综复杂，但都可以归结为"转化"和"构造"这两个范畴，其他的数学思想方法则处于更低层次的、从属的地位。所以，全书就围绕"构造"和"转化"这两条主线来展开讨论。

在这种思想指导之下，我们把全书分作六章：第一章概略地对我们提出的三类数学思想进行了宏观的扫描。第二章简要地介绍了数学史上一些重大的数学思想方法的产生和它们对数学发展的影响，希望能起到"温故而知新"的作用。第三章论述为什么要把"构造"和"转化"作为两大基本数学思想（指第三层

次数学思想或解题思想)，分析它们的意义、方法和作用．第四章讲述一些具体的构造思想．第五、第六两章讲述具体的转化思想．这里我们又把“转化”思想分成两类，第一类的转化比较自然，转化前后具有明显的、必然的逻辑联系，相当于我们通常说的解“常规题”；第二类的转化跳跃性大，转化前后缺乏明显的必然的逻辑联系，相当于我们通常说的解“非常规题”．

本书是在既分工又合作的情况下写成的．搜集材料、确定章节、提出论点、统一说法等等都是在共同讨论下完成的．至于执笔成文则分工进行．其中第一和第六章由欧阳维诚执笔．第二章和第五章由张垚执笔，第三章和第四章由肖果能执笔．作者的署名也是按所写章节的顺序安排的．

我们深深感谢本书的责任编辑郑绍辉先生，他从提出选题到审读书稿，付出了艰辛的劳动，提出了许多中肯的、宝贵的意见，如果说，这本小书对读者有所帮助的话，郑先生的辛勤劳动是功不可没的．我们也非常感谢我们的学友张楚廷教授，他在百忙中审读了我们的写作提纲和目录，并为本书作序．

作　者
千禧之年元旦于长沙

目　录

序 …………………………………………………………… (1)
前言 ………………………………………………………… (1)
第一章　什么是数学思想 …………………………………… (1)
§1.1　三个层次的数学思想 ……………………………… (1)
一、数学是什么 ……………………………………… (2)
二、数学向何处去 …………………………………… (5)
三、数学怎样论证 …………………………………… (10)
§1.2　数学思想的主要特征 ……………………………… (25)
第二章　数学思想发展史概述 ……………………………… (33)
§2.1　古代数学思想的两大源泉
——《几何原本》与《九章算术》…………………… (33)
一、《几何原本》及其影响 …………………………… (33)
二、《九章算术》及其影响 …………………………… (40)
§2.2　非欧几何的诞生 …………………………………… (45)
一、非欧几何思想的产生 …………………………… (45)
二、非欧几何创立的历史意义 ……………………… (48)
§2.3　三次数学危机 ……………………………………… (52)
一、$\sqrt{2}$的发现与第一次数学危机 ……………………… (52)
二、无穷小与第二次数学危机 ……………………… (54)
三、罗素悖论与第三次数学危机 …………………… (57)
§2.4　近代和现代数学思想的几次重大变革 …………… (62)
一、变量思想进入数学——解析几何与微积分的产生

及其意义 ………………………………………………………… (62)
二、随机思想的产生——概率论的创立及其意义 ………… (65)
三、代数结构思想的确立——群论的诞生及其影响 ……… (68)
四、模糊思想的产生——模糊数学的建立及其意义 ……… (69)
五、非标准思想的创立——非标准分析的建立及其
意义 ………………………………………………………… (72)
§2.5 电子计算机对数学思想发展的影响 ………… (75)
一、促进了数学思想方法的普遍化 ………………………… (75)
二、形成了新的数学学科和新的数学理论 ………………… (76)
三、促进了数学思想方法的新发展 ………………………… (77)
第三章 数学中的两大基本思想 ………………………… (81)
§3.1 两种基本数学思想 ………………………… (81)
§3.2 转化思想概论 ………………………………… (85)
一、转化的作用 ……………………………………………… (85)
二、转化的方向 ……………………………………………… (89)
三、转化的途径 ……………………………………………… (97)
§3.3 构造思想概论 ………………………………… (100)
一、构造思想的意义及构造法的分类 ……………………… (100)
二、构造思想的发展和作用 ………………………………… (106)
第四章 构造思想 ………………………………………… (110)
§4.1 数学中的构造思想 ………………………… (110)
§4.2 数学解题中的构造思想 …………………… (128)
一、用构造方法解数学问题 ………………………………… (129)
二、在解题过程中构造辅助工具 …………………………… (150)
§4.3 算法化思想与机器证明 …………………… (173)
一、算法与算法化 …………………………………………… (173)
二、几何定理的机器证明 …………………………………… (180)

第五章　转化思想(上)……………………………………………………(183)
§5.1　特殊化思想 ……………………………………………………(183)
一、寻找特例……………………………………………………………(184)
二、分类思想……………………………………………………………(189)
三、分步思想……………………………………………………………(192)
§5.2　一般化思想 ……………………………………………………(206)
一、模型化思想 …………………………………………………………(207)
二、强化命题思想 ………………………………………………………(211)
§5.3　变换思想 ………………………………………………………(214)
一、数式变形思想 ………………………………………………………(215)
二、变量替换思想 ………………………………………………………(219)
三、几何变换思想 ………………………………………………………(223)
四、命题转化思想 ………………………………………………………(227)
第六章　转化思想(下)……………………………………………………(238)
§6.1　映射思想 ………………………………………………………(238)
一、映射 …………………………………………………………………(239)
二、配对思想……………………………………………………………(244)
三、划分思想……………………………………………………………(255)
四、赋值思想……………………………………………………………(265)
五、表示思想……………………………………………………………(274)
六、同态思想……………………………………………………………(281)
§6.2　数形结合思想 …………………………………………………(291)
一、数形转化……………………………………………………………(292)
二、以数表形和以形验数 ………………………………………………(295)
三、图论知识的简单运用 ………………………………………………(303)

第一章　什么是数学思想

顾名思义，数学思想应该是人们在建立数学理论或解决数学问题时所用到的一些思想．但是，数学思想从来就不是数学家的专利，也不是只有在数学领域中才能发挥它的作用．社会正在逐渐走向理性，分析手法逐渐走向定量化，数学几乎渗透到了各个科学领域，数学思想也随之而渗入各个思想领域；反过来，各个科学领域的思想又反作用于数学思想，形成你中有我、我中有你的局面．例如，数学思想促成了电子计算机的发明和发展，而计算机的发展，又促进了数学机械化思想的飞跃．数学机械化思想将与公理化思想一样，受到数学家同样的甚至更大的重视．因此，可以毫不夸张地说，数学思想几乎存在于人类的各种思维活动中，具有最大的广泛性和深刻性．

§1.1　三个层次的数学思想

长期以来，人们在使用"数学思想"这一概念时，似乎并没有明确的界定，常常自觉或不自觉地强调了某一方面．一般地说，数学思想似乎可以分成三类或者说分成三个层次：

第一类是对数学本质的认识方面的思想，它回答"数学是什么"的问题．

第二类是关于数学的地位、作用和发展方面的思想，它回答"数学向何处去"的问题．

第三类是解决具体的数学问题的思想，它回答"数学怎样

论证"的问题，它与逻辑、策略以及具体的数学内容相联系．

一、数学是什么

这是一个古老而又新鲜的课题．许多数学家都从不同的角度探讨了这一问题．

1．众说纷纭的答案

英国著名的哲学家兼数学家罗素曾经说过："现代数学的主要成就之一，在于我们已经发现了数学实际上是什么．"

罗素所发现的数学实际上是什么呢？他写道："数学是这样一门学科，在其中我们永远不会知道我们所讲的是什么，也不会知道我们所说的是不是真的．"罗素的这一论断支持了古希腊学者柏拉图的论点．

柏拉图说："数学是现实的核心．"柏拉图把理性世界与现实世界割裂开来，认为永恒不变的关系只能存在于理性世界之中．数学中的关系是不变的，数学与人对现实世界的知觉无关．因此，数学才构成了现实世界的核心．

而马克思主义哲学则恰恰相反，认为数学必须依赖现实世界．恩格斯在《反杜林论》一书中强调说："数学是研究现实世界中数量关系和空间形式的科学．"

然而数学家柯朗与罗宾斯在他们合著的《数学是什么》一书中写道："数学是什么？对于学者和门外汉，这个问题都不是靠哲学来回答的，而是靠从事数学的经验来回答的．"

但是，即使是从事数学的人用经验来回答这个问题，也仍然是众说纷纭的．

本世纪关于数学基础的三大流派都在回答什么是数学这样一个问题．

逻辑主义学派认为，数学是由逻辑派生出来的．数学不过

是逻辑的一个分支.

直觉主义学派认为，数学是建立在可信性和构造性程序上的科学.

形式主义学派认为，数学是一种纯粹的符号游戏，对这种游戏的惟一要求是它不导致矛盾.

上述的各种说法显然有其片面性，还是柯朗与罗宾斯说得好："有一种观点严重地威胁着科学的生命：认为数学是一门不知所云的学科，它只是从定义和公理出发推导出来的一系列结论，除了这些公理必须互相一致以外，完全出自数学家心灵的自由创造．如果这种意见是正确的话，那么数学就会只由定义、法则和三段论法组成，成为没有动机，没有目标的游戏．认为有智慧的人可以凭其爱好创造有意义的假设系统．这种概念只对了一半，只有在内在必要性的指导下，自由思想才有可能达到具有科学价值的结果."

也有一些数学家从数学的作用来回答这一问题．如

皮尔斯：数学是得出必要结论的科学.

美国国家研究委员会的一些专家则认为：数学科学是集严密性、逻辑性、精确性和创造力与想像力于一身的一门学问．这个领域已被称为模型的科学.

庞加莱：数学是给予不同东西以相同名称的技术.

上述几种定义，显然有点模棱两可和难以捉摸.

哈佛大学数学家哈里森提出了自己的见解："数学是关于秩序的科学——它的目的就在于探索、描述并理解隐藏在复杂现象背后的秩序．数学的首要工具就是使我们能够将这种秩序说清楚的一些概念．正因为数学家花费了多少个世纪寻求最有效的概念来描述这种秩序的奇异特点，所以他们的工作就能应用于外部世界，而这个现实世界真可谓集复杂情况之大成，其

中有着大量关于秩序的问题．”

应该怎样理解他所说的“秩序”和“概念”呢？如果他所说的“秩序”是先于人提出的概念而客观存在的，那么，哈里森的说法就是亚里士多德的所谓“实物属性”的翻版．如果“秩序”是依赖于概念而存在的，而概念却是与现实世界无关的，那么，哈里森的说法就会成为柏拉图“理念世界”的翻版．

2．数学是关于秩序的科学

综合各种观点，我们能否认为

(1)数学是关于秩序的科学．当人们把现实世界的某些事物或关系抽象为概念之后，在这些概念中就隐藏着大量的、不以人的意志为转移的秩序．数学是通过概念和符号用逻辑推理的形式来表述这些秩序的科学．

(2)数学的一切进展都不同程度地植根于实际的需要；数学不管它多么抽象，它在自然界中最终必能得到实际的应用．当然，要准确地预测一个数学领域在何时何地将会派上何种用场是不可能的，也是不必要的．

(3)数学一旦在实际需要的背景下被推动了，它自身就会获得一种巨大的能量自我推进，很快超越直接应用的界限．因此，某种数学在一定的阶段可以是由未加定义的词语和公理组成的抽象系统，不必和自然有任何联系，不是任何事物的模型，而仅仅是因为数学家的兴趣与能力产生出来的，数学家没有把他的结论必须应用到自然中去的先决条件．

但是，如果根据某些数学内容是在远离现实的基础上，从抽象到抽象地产生出来的，就怀疑这些内容永远不会有用，就必然会受到时间的嘲弄．例如，著名的英国数学家哈代曾经十分自信地断言，抽象的数论不会有什么用处．但是，不到40年，数论竟然与保密工作发生了联系，在军事、经济和国家安

全方面发挥了重要的作用.

二、数学向何处去

这个问题实质上是前一问题的延续. 主要有两个方面的问题: 一是纯粹数学与应用数学之间的关系; 一是纯数学本身发展的关系. 毫无疑问, 这两个问题都影响到数学发展的方向.

1. 纯数学与应用数学的良性循环

纯数学与应用数学之间的关系, 人类文明的发展史已经证明, 它们实际上形成了很好的良性循环. 人类从现实的应用(诸如测量、生产、占筮、赌博等等)中抽象出了数学的概念. 当这些概念一旦形成, 数学家们就不再认为原来的那些用以抽象出数学概念的事物对数学本身有任何实际的意义, 因而彻底地抛弃它们, 依靠本身的动力, 在不与数学以外的任何事物进一步接触中自我成长. 数学家可以在这个自给自足、自我封闭的天地里充分地发挥自己的想像力和创造力. 从这个意义上讲, 可以说数学的本质是自由. 然而应用数学家不必担心, 不管纯数学走得多远, 随着时间的推移, 总能为它们找到实际的应用.

不幸的是, 为坚持纯数学的抽象方法所必需的各种手段——语言、符号、公式、定理等等掩盖了纯数学与应用数学之间的本质关系, 其结果是加大了纯数学与应用数学之间的人为的隔阂, 造成了纯数学与应用数学之间表面上的疏远. 于是, 一方面, 纯粹数学家可能认为自己的工作越来越令人不能理解, 在那象牙塔的顶尖上感到"高处不胜寒". 另一方面, 那些被一大堆抽象的符号和公式所吓倒的人们, 越来越认为数学只是一些玩弄公理的游戏, 而对它所带来的深刻见解视而不见.

我们不能被这种表面的现象所迷惑, 应该努力去保持和发

展纯数学与应用数学之间的良性循环．纯粹数学家可以心安理得地从事自己的研究，不必急功近利地为寻求它的应用而削弱理论的研究．应用数学家可以满怀信心地到已有的数学成果中去寻找对发展自己所研究的学科有用的东西．那里有取之不尽的宝藏，比其他任何资源更为有效、更为廉价．纯粹数学和应用数学好像人的两只眼睛，人们可以闭上一只眼睛去搜寻各自的目标，有时两只眼睛同时睁开可以看得宽一些，有时闭上一只眼睛（例如射击）却可能看得更准．

因此，过分强调数学为现实服务而削弱对基础理论的研究是不恰当的，它有可能阻碍数学甚至整个科学技术的发展．同样的，轻视应用数学对纯数学发展的作用也是不恰当的．过分强调经世致用的中国古代数学发展的道路就说明了这一点．中国古代数学曾经取得过辉煌的成就，但却长期被当作一种应用技术而在纯理论的发展上没有得到应有的重视．正是这种轻视纯数学理论研究而只重视数学应用的思想，在一段时期内延缓了中国纯数学的发展．这不能不说是中国古代（特别是近代）虽然在技术上也有许多发明创造，但整个科学发展却相当落后的重要原因之一．著名的科学家爱因斯坦曾经说过："西方科学的发展是以两个伟大的成就为基础，那就是：希腊哲学家发明的形式逻辑体系（在欧几里得几何中），以及通过系统的实验发现有可能找出因果关系（在文艺复兴时期）．在我看来，中国的贤哲没有走上这两步，那是用不着惊奇的．令人惊奇的倒是这些在中国全部做出来了．"这就是说，在中国古代，理论科学（特别是纯数学）没有得到应有的重视，虽然也取得了许多技术上的成就，但那要付出更多的代价．

另一方面，经世致用的思想，虽然延缓了我国现代数学的发展，但它却孕育和发展了算法化、机械化的数学思想，对西

方现代数学的发展产生了巨大的影响．但这一点却没有得到西方数学史家的认识与重视，如克莱因的《数学思想史》，只看到西方数学形式化、公理化的成就，而根本不提中国古代数学思想对数学发展的作用．就连李约瑟这样的学者，虽然注意到了中国古代科学的巨大成就，但也未能指出我国古代数学思想对现代数学发展的影响．这都是非常令人遗憾的．

2．数学基础的大辩论

正当人们对以集合论为基础建立的数学体系的精确性和有效性赞叹不已的时候，出现了关于集合论的“悖论”，从而爆发了数学基础的所谓“第三次危机”．

罗素的“理发师悖论”动摇了集合论，从而动摇了当时的数学基础．因为罗素悖论只涉及到集合论中的几个最基本的概念和一个基本原则：“集合”、“元素”、“属于”和“概括原则”，并不牵涉到其他技术性问题．罗素悖论所导出的矛盾，清楚地揭示了以集合论为基础建立起来的数学体系也会包含矛盾，同时数学中所采用的逻辑也是有问题的，这使当时的数学界和逻辑学界都同时感到震惊，这就是数学上的“第三次危机”．

数学是一门演绎推理的科学．从形式上讲，数学的真理性是以两点为支柱的：一是公理本身的真理性，二是逻辑规则的有效性．非欧几何的诞生动摇了第一根支柱，使人们看到了公理真理性的相对性．集合论悖论的出现，又动摇了另一根支柱．希尔伯特沮丧地写道：“由于悖论的出现而造成的局面是难于忍受的．只要设想一下，每个人曾学过、教过并在数学中加以应用的定义和演绎方法，从来都认为是真理和必然的典范，而现在却导致了荒谬．如果连数学思维都是不可靠的，那么到哪里还能找到真理和必然性呢？”

面对如此严峻的局面，人们首先对集合论的建立进行了清

理和补数，策墨罗引进了公理化，建立了形式集合论，并经过弗兰克尔、冯·诺伊曼等人的改进，成为形式集合论的所谓"ZFC系统"．ZFC系统能为数学(如分析与几何)提供严格的基础，即ZFC的无矛盾性可以保证数学的无矛盾性．但是，ZFC系统本身的无矛盾性却没有证明．ZFC系统只是"赶走"了悖论，并没有消除悖论．所以，曾经兴高采烈地宣称数学的严格基础已经实现了的庞加莱评论说："为了防备狼，羊群是用篱笆圈起来了，但不知道圈内还有没有狼．"不过，还值得数学家们聊以自慰的是，ZFC系统从建立到今天，尚未发现任何矛盾．

ZFC系统的建立，使由集合论的悖论所引起的数学基础的危机得到了缓解，但远未消除人们的疑虑．数学的可靠基础究竟是什么？数学的命题在什么情况下才具有真理性？对这些问题的不同回答，形成了数学基础中的各种学派，引起了20世纪初期关于数学基础的一场大论战．在这场争论中，对后来数学基础的发展有较大影响的有三大学派，他们是逻辑主义学派、直觉主义学派和形式主义学派．

(1)逻辑主义

逻辑主义学派的代表人物是罗素和怀特海．他们的基本观点是，数学就是逻辑，两者之间没有界线．"逻辑即数学的青年时代，数学即逻辑的壮年时代，青年与壮年没有截然的分界线，故数学与逻辑亦然．"罗素等人试图由逻辑推出整个数学．因此，数学不需要任何自己的公理，数学的概念完全由逻辑的概念导出，数学的命题也由逻辑的命题出发用纯逻辑推演出来．逻辑主义学派遇到了不可克服的困难，数学无法还原为逻辑，即由逻辑推导不出数学来．

逻辑主义学派虽然没有成功，但他们试图由逻辑推导出数

学的努力却大大地促进数学基础特别是数理逻辑的发展.

(2)直觉主义

直觉主义的代表人物是布劳威尔. 直觉主义学派主张以"直觉上的可构造性"作为"可信性"的标准, 对全部已有数学进行彻底的审查和改造. 他们认为排中律不能普遍适用, 数学对象的存在只在于可构造, 从而否认了相当一大批有重要意义的数学成果, 未免有失偏颇. 但他们提出的能行性问题具有重大意义, 他们的可构造性对数理逻辑与计算技术的发展起了重大的作用.

(3)形式主义学派

形式主义学派的代表人物一般认为是希尔伯特, 虽然他的观点与形式主义还有一些区别, 并且有些观点还与逻辑主义、直觉主义有共同之处. 希尔伯特提出了一个纲领, 其主要观点是: 首先将数学理论形式化, 然后用有限的方法证明形式系统的无矛盾性. 这里所说的形式系统, 是实现了彻底的抽象化和符号化的系统. 在这种系统中, 概念成了毫无现实意义的各种符号, 命题成了一个个由抽象符号组成的公式, 推导也成了抽象公式的某种变形.

希尔伯特的纲领, 由于哥德尔在 1931 年证明的两个不完全性定理而宣告失败.

希尔伯特在提出和实施这个纲领的过程中, 首次明确区分了三种数学理论: 一是非形式化的数学理论; 二是将第一种数学理论形式化而构成的形式数学系统; 三是以形式数学系统为研究对象的理论, 即元数学或证明论. 希尔伯特的"形式化"纲领虽然失败了, 但他建立的元数学却得到了发展, 成为数学一个重要的新分支.

三大数学流派虽然对 20 世纪数学的发展作出了贡献, 但

都没有解决数学向何处去的问题，其共同的致命弱点就是否认数学是植根于现实世界的、归根结底要受到现实世界的检验.

三、数学怎样论证

前两节我们对数学思想进行了客观扫描，现在回到微观方面来．某一数学体系是怎样建立起来的？某一数学命题是如何证明的？这涉及到数学思想中另一个最重要的层面，即人们在解决具体的数学问题时所用到的思想．也就是我们通常所说的数学解题思想．本书从第三章起，主要讨论这类思想．在没有特别说明的情况下，都是在这一意义下使用数学思想这一名词的．

我国著名数学家徐利治先生提出了"宏观数学方法论"与"微观数学方法论"的划分，按照徐先生的意见，关于数学发展规律的研究，属于宏观的数学方法论，微观的数学方法论则研究数学中的发现、发明与创新等法则．

根据徐先生的这一划分，本节谈到的第一、第二类数学思想应该属于"宏观的数学方法论"，而第三类数学思想毫无疑问应属于"微观的数学方法论"．

即使在微观的数学思想方法方面，也是名目繁多．通常所见，如优化思想、随机思想、划分思想、集合思想、映射思想、函数思想、方程思想、逼近思想、递归思想…凡此种种，不一而足．但究其目的，还是为了最后解决问题．所以为了说话的方便，我们把第三类数学思想或微观的数学思想方法简称为数学解题思想．

1．数学解题思想

数学思想既然是人们在解决数学问题时所用到的思想，不可避免地必须与具体的数学内容相关．虽然两者并不完全相

同，但却有着不可分割的联系，前者是解决后者的指导思想，后者是运用前者的具体成果．因此，本书采用以数学问题为中心来阐述数学思想的方法，力戒纯抽象的议论．

我们用著名的“七桥问题”为例来说明什么是数学思想．

18世纪时，东普鲁士的哥尼斯堡市（今属立陶宛共和国）城内有一条大河，河中有两个小岛，全城被大河分割成A、B、C、D四块陆地，河上架有七座桥，把四块陆地像图1－1那样连接在一起．当时许多市民都热衷于解决如下的一个难题：

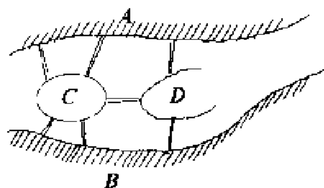


图1－1

一个散步者能否从某一块陆地出发，不重复地经过每座桥一次，最后回到原来的出发点．

这就是有名的哥尼斯堡七桥问题．

这个问题似乎不难解决，试验起来也比较容易，所以吸引了许多人都想来试试看，但是日复一日谁也没有成功．于是有人便写信向当时著名的数学家欧拉求教．欧拉毕竟是一位数学家，他并没有徒劳无益地再去重复人们已多次失败了的试验，而是首先产生了一种直觉的猜想：许多人千百次的失败，也许意味着这样的走法根本就不存在．于是欧拉把“七桥问题”进行了数学的抽象，用A、B、C、D四个点表示四块陆地，用两点间的一条连线表示连接两块陆地之间的一座桥，就得到了如图1－2那样的一个由一些点和点与点之间的一些连线组成的图形，这种图形在今天的数学中称为“图”．于是，所谓的“七桥问题”就转化为一个像图1－2那样的图是否可以“一笔画”的问题．

什么叫"一笔画"呢？那就是笔不准离开纸，一气画成一个图，但每一条线都只许画一次，不得重复．现在的问题是，像图1—2那样的图能否一笔画呢？1736年欧拉终于严格地证明了，答案是否定的．

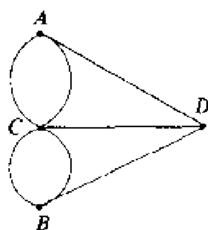


图1－2

问题的证明并不难理解，如果一个图可以一笔画的话，那么除了起点和终点之外，我们把其余的点称为中间点，对于每一个中间点都必须满足一个条件：与这个点相连的线必定有偶数条．因为对任何一个中间点来说，由于它不是起点，画笔必然沿某一条线到达此点；又由于它不是终点，画笔又必然要沿另一条线离开此点，这样进入这点几次，就要离开这点同样多次，一进一出，两两配对，所以与这个中间点相连的线必须是偶数条．这就证明了：一个可以一笔画的图最多只能有两个点，即起点和终点，可以由奇数条线相连(这样的点称为奇点)；其余的中间点都必须与偶数条线相连(称为偶点)．而图1－2中，A，B，D都与3条线相连，C则与5条线相连，四个点都是奇点，所以图1－2是不能一笔画的．从而证明了"七桥问题"所要求的走法是不存在的．

曾经难倒过许多人的"七桥问题"，就这样简单而圆满地解决了．但是，作为一个数学家，欧拉当然不会只满足于"七桥问题"的解决，他很快又进一步注意到："图"是有极丰富内容的数学对象，他敏感地意识到一门新的几何学即将诞生．这种几何学将不同于传统的欧几里得几何，它不考虑图形的形状和度量关系，只着眼于研究图形中点与线之间的位置关系，或者说它们之间相互连接的情况，于是在欧拉和其他一些数学家的共同努力下，一个新的数学分支建立起来了．欧拉当时称它为

"位置几何学"，现在则称为"图论"。

图论这一新的数学分支建立起来后，应用数学家很快地为它在生产、生活中找到了极为广泛的应用。因为许多问题所研究的对象，都是某一类特定的事物和这些事物之间的某种关系。如果用一些点来表示某一类特定事物的某些元素，用点与点之间的连线表示它们之间的某种关系，就得到一个图。换句话说，图可以作为许多实际问题的数学模型，从而人们可以借助于图论的方法去研究原来的实际问题。于是，人们最后抛开由点与线组成的几何图形，把图的概念进一步抽象化：

设 V 是一个有限集合，它的元素称为点；E 是 V 的点所组成的无序对所成的集合的一个子集，它的元素称为线，则二元组 V 和 E 一起称为一个图(这里的图指的是无向图)。

从"七桥问题"的试验到图论这一数学分支的建立，可以使我们从中清楚地看到数学思想运行的轨迹：

(1)首先，欧拉把一个实际问题通过抽象转化为一个数学问题。

(2)欧拉构造了图 1－2 那样的一个图，通过映射(陆地映射为点，桥映射为线)实现了(1)所要求的转化。

(3)问题转化为解决一笔画问题。欧拉利用逻辑分类的方法把图中的点分成奇点和偶点，问题再一次转化为讨论奇点的个数问题。

(4)由于奇点的个数非常明显，问题得以解决。但数学家并不就此满足，再一次把这个具体的图抽象化，转化为二元关系。

(5)再构造两个抽象的集合和它们之间的关系，给出图的一般定义，建立起"图论"这一新的数学分支。

从上面这个例子中，我们一方面看到了数学家锲而不舍、

不断创新的思想品质，另一方面也看到了数学家优秀的思维品质，这其中含有敏锐的直觉、合理的猜想、正确的解决、恰当的推广．他们解决问题的思想，主要是转化和构造．所以，我们认为数学思想的本质，在于转化和构造．转化是思维的进程，构造是实现的手段．不断地转化和构造，就成为解决数学问题的主线．所以，本书从第二章以后，主要围绕转化和构造这两大思想来展开讨论．

2．数学思想与数学方法

最后，我们谈谈究竟什么是“数学思想”？什么又是“数学方法”？两者之间有何联系与区别？

首先，我们认为：在本书开始提到的三类数学思想中的前两类，虽然对数学学科的发展具有指导性的作用，非常重要．但是在解决某一具体数学问题时，并不具有方法论的意义，所以，我们只把它称为“数学思想”，而不称为“数学方法”．

本书主要讨论第三类数学思想，它指解决具体数学问题时用到的一些思想，这些思想中又包含着方法．我们还是从具体问题入手进行讨论．

设 a，b，c 为 $\triangle ABC$ 的三边，S 为面积，求证：

$$a^2+b^2+c^2\geqslant 4\sqrt{3}S. \quad ①$$

本题是第3届国际中学生数学竞赛试题，数十年来人们对它一直很感兴趣，研究出了多种不同的解法，但笔者仍然欣赏下面这种解法：

首先考虑，当 $\triangle ABC$ 为正三角形的特殊情况时（如图1－3），我们有 $a=b=c$，$S=\frac{1}{2}ab\sin C=\frac{1}{2}a^2\sin 60^\circ=\frac{1}{4}\sqrt{3}\cdot a^2$，这时显然有

$$4\sqrt{3}S=3a^2=a^2+b^2+c^2.$$

即当$\triangle ABC$为正三角形时，①式是成立的(等号成立).

下一步我们设想将一般三角形向正三角形转化.

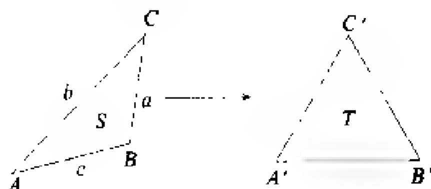


图1-3

再下一步设计出具体的转化方案. 构造一个边长为$\frac{a+b+c}{3}$的正三角形$A'B'C'$, 设其面积为T, 由等周定理, 知$S\leqslant T$.

要证①式，可转证加强的不等式：

$$4\sqrt{3}T\leqslant a^2+b^2+c^2. \quad ②$$

由于T已是可计算的，从而问题已获解决.

在本题的解法中，用到三层思想：

第一层次运用了从特殊到一般的思想. 它是解决任何问题都可能用到的思想，当然也是解数学问题要用到的思想.

第二层次运用了转化和构造的思想. 它设想把一般三角形的问题转化为特殊的正三角形处理，并构造出这样的正三角形.

第三层次用到等周化的思想. 根据它可以实际构造出正三角形$A'B'C'$, 更可具体计算出不等式②成立, 具有很强的操作性，而且与具体的数学内容的联系也更为紧密.

由此可见，我们谈的第二类数学思想也可以划分为三个不同的层次：

第一层次是广泛意义上的数学思想，它是人们在对待科学

研究、生产斗争，甚至立身处世、齐家治国都可能自觉或不自觉地用到的一些思想在数学中的表现．这种思想在数学中得到更清楚的表述和更有效的运用．如概念精确化的思想，量变与质变的思想、变化与不变的思想，整体与局部的思想，特殊与一般的思想，以退求进的思想等等．这一类思想虽非数学所专有，但在数学思想中经常地使用它们，不可或离，并且经过了数学化的表述，因此也是数学思想的一个层次．

第二层次是指那些与思维模式紧密联系的思想在数学中的表现．这类数学思想的一个重要特征是它不依赖具体的数学内容而存在，曾在各种逻辑思维方法中广泛地使用，但在数学中使其精确化、规范化，因而深深地打上了数学的烙印，充分地数学化而成为一种数学思想．如通常的抽象与概括的思想、归纳与演绎的思想、类比与猜想的思想、转化与构造的思想、公理化与模型化的思想等等，都是这一层次的数学思想．

第三层次是指人们在建立某一具体数学理论或解决某一具体数学问题时所用到的思想，这一层次的特点是它与相应的数学内容不可分离，在抽象的程度上还不能脱离具体的数学内容，不具有普遍性．如初等数学中的方程思想、函数思想，高等数学中的极限思想、群的思想、高维空间的思想等等，都应属于这一层次的数学思想．

上述数学思想的三个层次的划分只具有相对的意义，它们之间并没有明确的界线．

(1)在研究任何一个具体的数学问题时，三个层次的数学思想总是交叉使用的．第一层次的思想大抵相当于哲学思辨的层次；第二层次的思想则相当于逻辑分析层次；第三层次的思想则相当于知识运用的层次．如用解题作比喻，第一层次的数学思想大抵相当于通常所说的解题策略，第二层次的数学思想

大抵相当于解题方法，第三层次的数学思想则相当于解题工具．

（2）数学思想的层次性除了上述的三个层次之外，每一个层次还可分成若干层次．一个较低层次的数学思想经过适当的抽象就成为一个较高层次的数学思想．反过来，一个较高层次的数学思想具体地接触到某一特定的数学对象就转化为较低层次的数学思想．特别地，当某一数学思想已可对某些数学对象具体操作的时候，就称为"数学方法"．

（3）数学方法同样具有层次性．用"可操作性"来区分数学思想与数学方法，仍只具有相对的意义．一般地说，一种思想方法，抽象的程度愈高，就愈缺乏可操作性；抽象的程度愈低，就愈具有可操作性．一种数学思想，对于比其更抽象的数学思想而言可能是一种数学方法；相反对于比其抽象程度更低的数学方法来说，它又可能是数学思想．

因此，数学思想与数学方法很难作硬性的区分，在一般场合，统称为数学思想方法是适宜的．

例 1 解方程

$$\frac{x^4+3}{5}=\frac{3x^2+1}{2}.$$

通过去分母、移项，原方程化为

$$2x^4-15x^2+1=0.$$

再令 $x^2=y$，把双二次方程化为二次方程

$$2y^2-15y+1=0.$$

于是便可利用已熟知的二次方程求根公式来解了．

这个问题虽然很简单，但却交叉地使用了各个层次的数学思想．首先它运用了标准化与典型化的思想，因为解一元四次方程很不方便，但解一元二次方程却是很简单的，对于标准的一元二次方程有固定的解题公式．所以解题者首先想到把非标

准的对象转化成标准对象．这是第一层次的思想．其次它运用了转化思想，这是第二层次的思想．再次，还用了变量代换的思想，这是第三层次的思想．转化是一种数学思想，但具体到本题中，去分母、移项等是可以具体操作的．因而也是一种数学方法．又变量替换一般地说也是一种数学思想，但它较之转化这一数学思想显然抽象程度较低．而且在本题中，它具体地为转化思想服务，具有令 $y=x^2$ 这样的可操作性，所以又是一种数学方法．

例 2 设正整数 d 不等于 2，5，13．证明：在集合｛2，5，13，d｝中可以找到两个不同元素 a，b，使 $ab-1$ 不是完全平方数．

这是 1986 年于波兰华沙举行的第 27 届国际中学生数学竞赛试题．一种解法如下：

若 $2d-1$，$5d-1$，$13d-1$ 都是完全平方数，则必 $2d-1\equiv0,1(\mathrm{mod}4)$，$5d-1\equiv0,1(\mathrm{mod}4)$，$13d-1\equiv0,1(\mathrm{mod}4)$．

(1) 当 $d\equiv0,2(\mathrm{mod}4)$，则 $2d-1\equiv3(\mathrm{mod}4)$，矛盾．

(2) 当 $d\equiv1(\mathrm{mod}4)$，则令 $d=4h+1$．

若 $h\equiv1,2(\mathrm{mod}4)$，则 $5d-1=2^2(5h+1)$，而 $5h+1\equiv2,3(\mathrm{mod}4)$，矛盾．

若 $h\equiv0,3(\mathrm{mod}4)$，则 $13d-1=2^2(13h+3)$，而 $13h+3\equiv2,3(\mathrm{mod}4)$，矛盾．

(3) 当 $d\equiv3(\mathrm{mod}4)$，则 $5d-1\equiv2(\mathrm{mod}4)$，矛盾．

综上所述，不论 d 为何种正整数，$2d-1$，$5d-1$，$13d-1$ 中至少有一个非平方数，从而命题得证．

在本题的解法中，也交叉地使用各个层次的数学思想．首先，它运用了正面和反面、特殊与一般等第一层次的数学思想．其次它运用了逻辑分类、穷举归纳、转化(把判断平方数

转化对模 4 的余数的讨论)、分步(对 $d\equiv1\pmod4$ 再分类)等第二层次的数学思想. 再次, 它使用了同余等第三层次的数学思想. 同余本是一种重要的数学思想, 但在本题中它作为实现"分类"这种更抽象的数学思想的工具, 在模 4 上具体操作, 又变成一种数学方法.

3. 两种基本思想——构造和转化

数学解题思想如此纷繁, 我们却选取转化思想和构造思想作为两大基本思想, 这样做的理由在哪里呢?

首先, 从字面上讲, 转化就是把一种形式变成另一种形式. 解数学题, 其内容虽然千差万别, 但就其形式看, 无非是从条件出发, 逐步向结论靠近, 每靠近一步, 都是一种转化, 可以毫不夸张地说, 转化能力的高低是衡量数学家解题能力的重要标志.

从某种意义上讲, 许多数学思想都可归结为转化思想的一种或作为实现转化的一种手段. 例如, 变量替换是一种重要的数学思想. 欲解双二次方程

$$x^4-6x^2+5=0,$$

我们利用变量替换的思想, 令 $x^2=y$, 则这个双二次方程转化为:

$$y^2-6y+5=0\ .$$

运用变量替换这一数学的思想的目的, 是为了完成从双二次方程向二次方程的转化. 在这里转化是最基本的思想, 变量替换只是为完成这一转化而用到的较低层次的思想, 它可以归结为转化思想的一种手段.

又例如, 我们证明余弦定理:

在$\triangle ABC$ 中, 设 a、b, c 分别为$\angle A$、$\angle B$、$\angle C$ 的对边, 则有

$$c^2 = a^2 + b^2 - 2ab\cos C.$$

在论证这一定理时，我们运用了分类的思想，按 C 为直角、锐角、钝角来分别讨论。但是从根本上看，这是一种转化，把一个整体的问题转化为三个较小的问题，以便分而治之，各个击破。在这里，分类思想只是实现转化的方法。

在数学中论证问题一般采用演绎法与分析法。按照演绎法的推理过程：

若有 A，必有 B；

若有 B，必有 C；

…

若有 Y，必有 Z。

从条件 A 到结论 Z 表现为命题：

若有 A，则有 Z。

在演绎过程中，每前进一步（不管它是通过什么数学思想方法来实现的），都实际上完成了一种转化，即把命题

若有 A，必有 Z

的形式，转化为命题

若有 B，必有 Z

的形式。通过不断地转化，逐步将条件 A 与结论 Z 联系起来，获得了问题的解决。

因此，把转化思想作为一种基本的数学思想是合理的、必要的。

但是，转化并不总是永远畅通无阻的。正像走路一样，我们要迈向目的地——结论，从条件出发，不断地从一处转向另一处，逐渐向结论靠拢。但有的地方却无法过去，需要修筑道路，架设桥梁，这就需要构造。例如，我们要证明几何命题：

等腰梯形的底角相等。

在等腰梯形 $ABCD$ 中，$AB=DC$，$AD\parallel BC$．求证：$\angle B=\angle C$（如图 1－4）．

欲证两角相等，可转化证两个全等三角形的对应角相等．但是，现在并不存在两个现成的全等三角形以供转化．所以，过 A 作辅助线 $AE\perp BC$，过 D 作 $DF\perp BC$，构造出两个全等的直角三角形，下一步就可转化为论证两个三角形全等了．

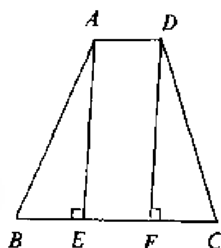


图 1－4

又例如，组合数学中有一个著名的拉姆赛数问题：

平面上有 n 个点，将它们两两之间连一条线，称为完全 n 边形．现在将一个完全 n 边形的所有线都染成黑、白两种颜色，如果能保证不管怎样染色，一定会出现一个完全由黑色线段组成的完全 p 边形，或者出现一个完全由白色线段组成的完全 q 边形．满足这一条件的最小正整数 n 称为拉姆赛数，记作

$$R(p,\ q)=n\ .$$

已经知道 $R(3,\ 3)=6$．要证明这一结论必须分两个方面．一方面证明，在 6 点之间两两连一条黑线或白线，至少会出现一个黑色的三角形或一个白色的三角形．另一方面，必须证明在 5 点之间两两连一条黑线或白线，可能出现既没有白色线段组成的三角形，也没有黑色线段组成的三角形的情况．为了证明后者，我们很难把它转化为另一命题，必须构造出这样的一种连线法，使它的确不出现同色的三角形．只要构造出图1－5，就完成了我们的证明．

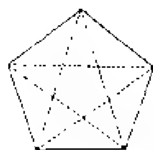


图 1－5

由此可见，转化和构造是数学思想中的两个基本思想，其他的一些思想从某种角度看可以归结到这两种基本思想中来．因此，解数学题的过程，可以看作是转化与构造交互使用，逐步推进的过程．

4．数学思想方法的分类

具体的数学思想，名目繁多．通常所见，有逻辑思想、非逻辑思想、构造思想、转化思想、公理化思想、数学模型思想、优化思想、随机思想、划分思想、集合思想、映射思想、函数思想、方程思想、数形结合思想、逐次逼近思想、极限思想、递归思想、一般化思想、特殊化思想、演绎思想、归纳思想…凡此种种，不一而足！

数学思想方法如此纷繁，当然给研究带来困难，因此有必要对它进行适当的分类．要对数学思想方法进行分类，首先得有一个分类的标准．例如，可以从它们的层次性、功能性、学科属性等方面提出分类标准．究竟应该采用什么标准分类，在目前所见的文献中似乎还缺少这方面的研究工作．我们认为：关于数学思想方法分类的综合研究，还有赖于对各种数学思想方法及其相互关系的研究工作的积累和深入．在没有提出对繁多的数学思想进行分类的合适的标准之前，选择若干能够体现数学和数学思想的本质特征的思想，用以统领和概括其他各种数学思想，组织成一定的体系，应该是一种可行的研究方法．

我们在前面已经把“构造”和“转化”作为两个基本的数学思想方法，那是从层次和功能两个方面着眼的．

首先不难发现，上面所列的种种数学思想并非完全处于同一个层次上，逻辑思想、非逻辑思想、构造思想、转化思想等的层次较高，而方程思想、函数思想、极限思想的层次较低．如函数思想可归入映射思想，而极限思想则可归入逐次逼近思

想，而映射思想和逐次逼近思想最后亦将归入转化或构造的思想，从这个意义上讲，似乎应该按层次性对数学思想方法进行分类。

但是，另一方面，单纯地按层次性进行分类，有时并不方便。

例 甲、乙两队各出 7 名队员按事先排好的顺序出场参加围棋擂台赛，双方先由 1 号队员比赛，负者被淘汰，胜者再与负方 2 号队员比赛，…直到有一方队员全被淘汰为止，另一方获得胜利，形成一种比赛过程。那么所有可能出现的比赛过程有多少种？（1988 年全国数学竞赛高中联赛试题）

解法一：我们只要考虑甲队获胜的情况就可以了，因为由于对称性，乙队获胜的情况与甲队获胜的情况是一样多的，我们只要能求出甲队获胜比赛过程的种数，再乘以 2，就得到全部比赛过程的种数。

设甲队队员出场顺序为 A_1、A_2…、A_7，A_i 获胜的局数为 x_i，$0\leqslant x_i\leqslant 7$，$1\leqslant i\leqslant 7$，则有方程

$$x_1+x_2+\cdots+x_7=7. \qquad ①$$

易知，甲方获胜的比赛场次的集合与不定方程①的非负整数解集之间可建立一一对应的关系。例如，不定方程①的非负整数解

(2，0，0，1，3，1，0)

就对应这样一种比赛过程：

A_1 胜两局，即击败了 B_1、B_2 但负于 B_3；B_3 连胜三局，即胜了 A_1、A_2，A_3 但负于 A_4；A_4 仅胜 B_3 一局即负于 B_4；A_5 连胜 B_4，B_5，B_6 但负于 B_7；A_6 战胜 B_7，比赛结束。

可知，甲队获胜的不同比赛过程总数是方程①的非负整数

解的个数 C_{13}^{7}.

解法二：考虑由 7 个 B 和 6 个 A 排成的序列，例如：

$$AABBBABBABAAB \qquad ②$$

序列②对应这样一场由甲队获胜的比赛：第 1，2 局甲队的 A_1，A_2 被淘汰，第 3，4，5 局乙队的 B_1，B_2，B_3 被淘汰；第 6 局 A_3 被淘汰；第 7，8 局 B_4，B_5 被淘汰；第 9 局 A_4 被淘汰；第 10 局 B_6 被淘汰；第 11，12 局 A_5，A_6 被淘汰；第 13 局 B_7 被淘汰，比赛结束，甲队获胜.

甲队获胜的比赛场次与排列②之间可建立一一对应关系. 排列②由 13 个位置中 6 个 A 的位置决定，因此有 $C_{13}^{6}=C_{13}^{7}$ 种，即甲队获胜的比赛过程有 C_{13}^{7}种.

对同一个问题的两种解答都用到了映射的思想，但是，它们在两种解法中不仅层次不同，而且功能也有差别.

在解法一中，首先只是通过构造的方程建立一个不定方程②，在构造出的方程中，其非负解与比赛过程可建立对应关系. 在这里映射思想是隐含于构造之内的，即构造出来的方程能发挥映射的功能，而不是通过映射来构造出一个方程. 因此，它属于较低的层次.

而解法二却不是这样，它是直截了当地把比赛过程与排列②对应，通过较易计算的排列数求出比赛过程的种数，是一种直接的转化思想，而且也属于较高的层次.

由此可见，即使是在解决同一个问题中用到的同一种数学思想方法，它们既可以有不同的层次，也可以有不同的功能.

因此，对数学思想方法的分类，应该兼顾它的功能和层次. 首先可将各种各样的数学思想按其主要功能分类，主要功能相同或相近的归入一类，在同类中再比较其层次. 例如“变换思想”是把一个数学的表述从一种形式变换为另一种形式，

即用不同的数学语言表达同一个问题．如同普通语言中的“翻译”一样，是把一种语言翻译为另一种语言．而“数式变形”思想，如解方程的同解变形，证恒等式的恒等变形等等；“换元思想”，如三角代换等．都是一种变换思想．又在“数式变形”中，我们常使用“配凑思想”等．“变换思想”、“数式变形思想”、“换元思想”、“配凑思想”都属于“变换思想”（就其功能而言），也同属于转化思想．但它们的层次是不一样的，除了最高层次的转化思想以外，“变换”的层次较高，“数式变形”和“换元”层次较低，它们可归结为“变换”的一种；而“配凑”的层次更低一些，而且可以归结为“数式变形”的一种，等等．

另外，我们把数学解题的过程看作是一个不知解法的问题向已知解法的问题逐步靠拢的过程，把每一步靠拢都看成一次转化，其程式是：

$$A \xrightarrow{\text{转化}} B \xrightarrow{\text{转化}} C \cdots \xrightarrow{\text{转化}} Z$$

（求解的问题）（中间问题）（已知解法的问题）

如果这些中间问题 B，C，…都是不专门为 A 而构造出来的现成命题，那就只有一系列的转化过程．如果某一中间问题是依赖 A 才能够构造出来的，这就涉及到“构造”问题．

当然，必须指出的是：我们在这里提出的分类方法，只是为了本书展开讨论的方便．如前所述，真正合理的分类方法还有待进一步深入研究．

§1.2　数学思想的主要特征

数学思想作为一种科学思想，除了具有一般科学思想的特征外，还具有其独特的性质．通过对这些独特性质的讨论，将有助于我们更深刻地理解数学思想．

在这一节里，我们将粗略地讨论一下数学思想的几个主要特征：抽象性、自由性、辩证性、一致性和包容性。

1. 抽象性

一提到数学的特点，人们就常用高度的抽象性、严密的逻辑性、广泛的适用性来加以概括。如前所述，数学是研究从客观世界抽象出的概念中所隐含的普遍秩序的科学，客观世界的现象是复杂的，越是当数学从现实出发，退到极端抽象的区域时，它就越能在分析具体事实方面得到脚踏实地的应用。换句话说，正是这种扎根于现实的极端的抽象使我们能用以作为控制具体事实的思想武器。因此，在建立数学概念时进行了高度的抽象，不断地舍弃了事物的更多具体性质，自然就带来概念适用的广泛性。又因为概念的高度抽象，为了保证从概念出发建立起来的理论体系的正确性，推理只能采用演绎的方法，也就是要求严密的逻辑性。因此可以说，高度的抽象性是数学概念和数学理论最重要的特征，而用以建立数学概念和数学理论的数学思想也是以高度的抽象性为其主要的特征的。

任何科学的概念都是对被研究对象进行抽象的结果，但抽象的程度是不同的。自然科学的许多概念，虽然也非常抽象，但都在现实世界中具有直接的原型。例如，物理学中的"量子"概念，化学中的"化学键"的概念，都是很抽象的，但是这些概念都可以在客观世界中找到现实的原型。数学的概念就不同了，除了像自然数、三角形等极少数的原始概念是从现实世界中抽象出来的以外，大多数数学概念，诸如复数、向量、线性空间、群、泛函等等，都是在"远离现实"的基础上定义出来的。

即使是最原始的概念之一的自然数，它是由人们的计数活动中产生的，但却不是由被计数的客观存在的事物抽象出来

的．例如．"1"并不是从1个指头，1只苹果等具体事物中直接抽象出来的．它建立在所谓"等置抽象"的基础上，反映的是集合的类的特征．我们不难验证，一一对应关系是一种等价关系，按照这一关系把事物进行分类，把各种彼此间具有一一对应关系的集合归为一类．自然数(在基数意义下的自然数)就是同一类中的所有集合的共同属性．

2．自由性

数学思想的本质是自由．这里所说的"自由"不是说数学思想可以随心所欲地胡思乱想，也不是说数学思想可以永远不接受客观实践的检验．恰恰相反，所有的数学思想，始终扎根于现实并最终要接受客观的检验．这里所说的数学思想的自由性是指：

数学思想可以冲破各种束缚，可以冲破已有的经验和认识的定论，可以不必每前进一步都必须接受客观现实的检验(当然最终还是要接受客观现实的检验)．数学家最能打破传统的、僵化的思想．一个数学家的座右铭是"怀疑明显的东西"．在其他科学中，例如物理学，所有的理论都必须用实验来证实，才能最后肯定其正确性，数学思想则不必这样循规蹈矩．

数学思想不但需要经验和推理，更需要直觉与想像．数学思想所表现出来的创造性和想像力是其他科学很难赶上或超越的．

数学思想产生的动机也是比较自由的，很少受到功利思想的约束，在研究问题上有较大的选择自由度，哥德巴赫猜想的研究、费尔马大定理的研究，很难说有什么明确的动机或功利的目的，只是少数数学家的兴趣和爱好．但正是在这些研究中，创造了极其重要的、有价值的数学思想．

罗巴切夫斯基创立非欧几何的过程最能说明数学思想的自

由性．欧几里得几何学经过了两千多年的历史检验，奠定了绝对正确的权威，"三角形三内角之和等于180°"是放之四海而皆准的真理．但罗巴切夫斯基却不顾这些而提出怀疑，并凭自己丰富的想像力建立起自己的一套几何学．尽管在长时间内，他的几何学既找不到现实的模型，也得不到社会的承认，但他照行不误，不断地追求下去，终于获得了成功．

3．辩证性

数学思想的另一重要特征是它的辩证性．这是因为客观世界的本身就是各种各样的、复杂的、充满着辩证性的．数学最终是反映客观现实的本质属性的，也就必然充满着辩证性．所以数学思想充满着辩证思想．例如矛盾的观点、变化与不变的观点、统一性与特殊性的观点、总体与局部的观点等等．

数学思想注重事物的统一性和特殊性．

微分中值定理由三个定理组成：罗尔(Rolle)定理，拉格朗日(Lagrange)定理及柯西(Cauchy)定理．这三个定理可以认为同时立足于这样一个几何事实：任何光滑的曲线段至少存在一点，在该点的切线平行于连接曲线两端点的弦．罗尔定理是拉格朗日定理的特殊情况，而拉格朗日定理又是柯西定理的特殊情况．但是这三个定理又有一个统一的表达：设函数 $f(x)$、$g(x)$、$h(x)$ 在闭区间 $[a, b]$ 上连续，在开区间 (a, b) 内可导，那么在区间 (a, b) 内至少有一个点 ξ，使得

$$\begin{vmatrix} f(a) & g(a) & h(a) \\ f(b) & g(b) & h(b) \\ f'(\xi) & g'(\xi) & h'(\xi) \end{vmatrix} = 0 .$$

显然，当 $h(x)=1$，$g(x)=x$，且 $f(a)=f(b)$ 时是罗尔定理；当 $h(x)=1$，$g(x)=x$ 时是拉格朗日定理；当 $h(x)=1$，且 $f(a)\neq f(b)$，$g(a)\neq g(b)$ 至少有一个成

立时是柯西定理.

数学思想重视事物的总体性.

三角形的内角和等于180°.这是一个简单的数学命题.但从某种角度看,它又是一个深刻的命题.它的深刻之处是因为它阐述的是三角形的总体性质.三角形的内角和是180°,四边形呢?五边形呢?容易证明:四边形的内角和等于360°,五边形的内角和是540°.一般地n边形的内角和是$(n-2)$个180°.

这个公式好像给出了一般规律,其实不然.把视角改变一下,不看内角而看外角,则有

三角形的外角和为360°;

四边形的外角和为360°;

五边形的外角和为360°;

…

n边形的外角和为360°.

这个结论就比"n边形的内角和是$(n-2)180°$"更为简明而漂亮,从而实现了从特殊到一般的转化.

多边形是由有限条线段组成的,当有限转化为无限时,多边形就转化为一条封闭曲线.质点沿一条不自相交的封闭曲线运转一圈,运动方向的改变量的代数和恰好是360°.这里又一次从特殊达到了一般.

数学思想关心变化中的不变量.

几何学的任务就是研究某种特定变换群下的不变性质,换言之,不同的几何学就是研究在不同的变换群下不变性质和不变量的科学.

运动变换(平移和旋转)改变图形的位置,但不改变图形的形状和大小,线段的长度即两点间的距离不变.

相似变换改变了图形的大小，但形状不变．即两点间的距离变了，但保持直线变为直线，直线之间的角度也不变．

仿射变换改变图形的形状和大小，但仍保持直线变为直线，平行线变成平行线．

射影变换可能把平行线变成不平行线，但仍保持把直线变为直线．

拓扑变换不再保持上述各种不变性质，但仍保持某些不变性．例如，一个圈总变为一个圈．在拓扑学中还有一个布劳威尔不动点定理："若 A 和闭单位圆盘 D 同胚，f 是 A 到 A 的连续映射，则 f 在 A 上必有不动点．"

各种几何变换都有不变的东西，从而构成相应的几何学．

数学思想中充满着辩证性，无怪乎在中学数学教育中，把培养学生的辩证唯物主义观点作为一项重要的教育目的．

4．一致性

数学思想具有惊人的一致性．

俗话说："人同此心，物同此理"．这一说法在数学思想中得到了最确切的证实．在许多科学领域中，对于同一事物，不同的学者往往得出大相径庭的结论．对于数学则不然，对于同一问题，人们可能用不同的方法研究，不同的方式思考，但总是殊途同归，得出同一结论．思维路线存在着同构和平行．

数学思想的一致性是由它的本质决定的．数学由于它高度的抽象性和表现的自由性，造成了数学思想在发展道路和表现形式的差异，但最终又是反映客观世界最本质最集中的属性的，所以这些形异实同的思想又必须趋于一致．

本世纪关于数学基础的三大流派，尽管他们开始时各执一端，互相激烈的批判，但到最后又不得不互相吸收对方的观点来支持自己的理论．

勾股定理据说有几百种证法．毕达哥拉斯学派从抽象的演绎证明了它；中国的古代数学家则用实际的计算证明了它．但两者的思想是一致的，都是利用面积割补或比例线段．无怪乎有数学家建议，如果我们将来要与外星人交流而语言不通时，就可用勾股定理来和他们对话．因为无论是地球人或者外星人，他们对数量的理解总是一致的．这也许是一句玩笑话，但却从侧面说明了人类数学思想的一致性．不知道九重天外，是否真有知音？

5．包容性

在科学领域中，两种互相对立或矛盾的思想，一般是不能长期共存的．两者谁是谁非，孰真孰假，终将受到实践效果的检验．新的思想或者在旧思想面前难以立足，落荒而逃；或者推翻旧的思想，取而代之．如伽利略在比萨斜塔上的著名实验，就彻底推翻了亚里士多德的"重物先落地"的结论．在数学中则不一定是这样，两种完全对立或矛盾的思想，却可以和平共处，求同存异，最后达到和谐的统一．

非欧几何与传统的欧几里得几何是矛盾的，但它们都能在数学的大厦中各自安家立户，最富有戏剧性的是，第一个罗氏几何的数学模型是在欧氏几何中找到的．欧氏几何用自己的事实支持了自己的对立面．争论它们的谁是谁非，曾经是一场令人遗憾的、毫无意义的误会．

数学思想的包容性也是由数学的本质所决定的．数学既然最终是反映客观世界最普遍的秩序的，数学中一切有生命、有价值的成果都是基于不同的领域的事实与思想的几经曲折、几经抽象的反映，它们虽然千姿百态，甚至互相矛盾对立，但必须看到他们的整体性．数学就像一片无限广袤、风景优美的原野，数学思想只不过是在这广阔天地中从不同角度挑选出的几

个景点，景外有景，天外有天，横看成岭，侧看成峰，智者乐山，仁者乐水，决不可以把自己立足的、看到的某一景点当作整个天地的全貌。

第二章　数学思想发展史概述

数学的产生和发展经历了漫长的历史过程，到现在至少有五千多年的历史了．各文明古国在发展数学上都做出了自己的贡献．温故而知新，我们现在研究数学思想，回顾前人走过的路程，从中汲取有益的经验，是完全必要的．因此，在具体讨论数学思想的若干问题之前，我们选择一些重要的、对数学思想的发展有代表性的数学史实，有重大影响的成果略加叙述，以揭示数学思想发展的历史概貌以及与我们今天所讨论的数学思想的继承关系．

§2.1　古代数学思想的两大源泉——《几何原本》与《九章算术》

在各种古代数学思想中，其体系性较强且对现代数学思想的形成和发展影响较大的要数古希腊数学和中国古代数学．这两种古代数学的历史非常悠久，其发展都在千年以上，内容十分丰富．本节中不准备作全面探讨，只对其中最具有代表性的两部著作——古希腊的《几何原本》和中国古代数学名著《九章算术》进行初步的研究，着重探讨它们形成的历史背景、思想特点及其对现代数学思想发展的影响．

一、《几何原本》及其影响

1.《几何原本》形成的历史背景

《几何原本》是古希腊数学家欧几里得的著名著作．它是古

希腊数学的系统化总结，它的产生有着深刻的历史背景。

公元前7—公元4世纪，已是希腊奴隶制度发达时期，埃及的几何学及巴比伦的算术先后流入希腊。一方面，手工业、商业和造船业的发展需要力学、天文学及航海等方面的知识；另一方面，由于社会物质的增加，使得一部分奴隶主阶级中的成员可以专门从事脑力劳动而不从事物质生产的劳动。这就使得作为奴隶主阶级中成员的古希腊数学家十分鄙视"应用"数学。古希腊数学家建立的数学体系中之所以不包括应用数学的内容，这是一个很重要的原因。

古希腊是一个城邦国家。希腊的哲学家(也是数学家)都致力于辩论术的研究，而数学是辩论有力的工具，且辩论所需的数学与农业、手工业及商业中需要的数学是不相同的，辩论中需要概念准确，推理严谨，还需要有随机应变的各种方法和技巧，因而古希腊的数学发展是与形式逻辑的发展紧密相关的。把形式逻辑的思想方法运用于数学研究和排斥数学应用成为当时一股强大的思潮，这种思潮在一定程度上左右了数学发展的方向。于是，发展了公理化的方法，强调抽象化的理论，形成了封闭的演绎体系。

早期游历埃及、巴比伦的希腊学者泰勒斯首先开始采用数学证明方法，后来，毕达哥拉斯及其学派的成员企图用数解释一切，不仅认为万物都包含数，而且说万物都是数，他们发现了包含勾股定理(国外称为毕达哥拉斯定理)在内的一系列几何命题。由此，数学由具体的、经验的阶段过渡到抽象的、理论的阶段。爱尼亚(Elea，意大利半岛南端)学派的巴门尼德第一个采用了反证法，这一学派的芝诺在证明中用了归谬法，而且他提出了四个悖论(二分悖论、追龟悖论、飞箭不动、运动场悖论)，对逻辑的发展起到巨大的推动作用。雅典学派的柏拉

图首先把形式逻辑用于数学，他要求数学定义要准确，假设要清楚，并阐述了许多论证方法、定义方法及一些逻辑规律和要求．他的学生亚里斯多德更加努力把形式逻辑应用于数学，他首先指出一个数学体系中的概念必须用已知概念来定义，并由此指出数学体系中必须存在不加定义的概念．他和柏拉图一样，认为应当把一些命题承认为'理所当然'的假设，"凭此假设为起点推演他们自己的论题"，还说"在几何上有些命题不证而明，而其他一切命题或多数命题的证明却依赖于这些命题，我们称这些命题为几何的要素"．他强调数学家必须这样做，把这些假设称为公理或通则，用它们作为推理的基础，对于"那些专门进行研究的人——如几何学家或算术家——均不问这些通则是真是假"，并在实际上给出了后来欧几里得所采用的某些公理．此外，他对各种逻辑规律，例如矛盾律、排中律、同一律都作了论述，指出他们在逻辑证明中的重要意义，他还发展了三段论法，实际上，他的三段论体系就是一个初级的公理体系．总之，亚里斯多德奠定了古希腊数学的逻辑基础，但他没有实际用过公理化方法去推证定理，也没有构造出一个公理化的知识体系．

反对数学用于实践的思潮及形式逻辑的发展对古希腊的数学产生了深远的影响，《几何原本》正是在这种背景下出现的，可以说是当时特定的学术思潮下的必然产物．

2.《几何原本》的主要内容及思想特点

《几何原本》使得数学成为了一门独立的、具有演绎系统的科学，对东、西方数学思想的发展产生了深远的影响．

《几何原本》共 13 篇，第一篇至第四篇是关于平面几何——直线形和圆的理论，第五篇是比例论，第六篇是平面相似形，第八、九、十篇则论述算术(数论)，第十篇是关于"不可

通约量"的理论，第十一、十二、十三篇是关于立体几何的理论和"穷竭法"．从内容来看，《几何原本》包括了目前中学课程里的平面几何、立体几何、初等数论的内容，它不仅在传播几何知识方面发挥了巨大的作用，而且在推进数学思想方法向前发展方面产生了深远的历史影响．

《几何原本》在数学思想方法方面有如下的特点．

(1)采用了公设化的叙述方法

《几何原本》全书有 475 个(有的版本有 477 个)命题，首先给出 5 个公理和 5 个公设(公理在所有学科中都适用，而公设只适用于几何学，现在人们不加区分，一律称公理)，然后给出 23 个定义，其余 465 个命题都是由这 10 个公理或公设及相关概念的定义运用三段论推导出来的，除了个别定理证明不够严格(例如，利用了图形的直观)以及个别定理证明有错误外，绝大多数证明用今天的观点来看也是正确的．虽然，亚里斯多德制定了科学演绎理论的方法论原理，但是，他对公理化思想方法的实际功能估计不足．他曾经认为，演绎的"出发点在数目上比结论少不了多少"．《几何原本》用少量的公理推出大量定理的这一事实说明，公理化思想方法具有极其重要的科学方法论功能．在一个演绎体系中，作为推理的出发点的公理和概念，并不一定需要很多，相反，需要选得尽可能少．《几何原本》成为了用公理化思想方法叙述数学内容的典范．对现代数学及一些科学的发展产生重大影响的公理化思想正是起始于欧几里得的《几何原本》．

(2)建立了比较严密的演绎系统的逻辑结构

《几何原本》是用公理化思想方法表述的．全书从少数几个不加证明的命题(公理和公设)出发，按照一定的逻辑规则，演绎出其他所有的命题来．除了推导所需要的逻辑规则外，《几

何原本》是由一系列公理、定义、定理等构成的数学理论体系．在《几何原本》中，定理的排列顺序是经过认真考虑的，每一个定理的证明允许采用的论据只有公理和前面已证过的定理，原则上不依赖别的其他东西，当然，在个别命题上违背了这一原则，例如，对《几何原本》中第一卷第四命题："若一个三角形两边及其夹角相应地等于另一个三角形的两边及夹角，则两个三角形相等"，欧几里得采用的证法是将一个三角形重合到另一个三角形上去．然而，要使两个三角形重合，必须借助于运动，还必须假定运动中图形的长度、角度均保持不变．《几何原本》中却没有相应的公理，于是，证明中不得不借助于图形的直观．后来，许多数学家正是在《几何原本》的这种数学思想指导下，不断排除体系中的直观部分，逐步得到了更严格的数学演绎体系．

（3）重视抽象化理论而忽视应用

《几何原本》中研究的是一般的、抽象的概念和命题以及概念、命题之间的逻辑关系，即如何由已有的一些概念和命题定义出另一些概念和推演出其他一些命题，而不考虑产生这些数学概念和命题的实际背景，也不研究这些数学概念和命题的实际原型．例如，在《几何原本》中研究了抽象的"数"的性质，却不涉及到具体的数的计算与应用．它排斥各种理论的实际应用，全书四百多个命题中没有一个是研究实际应用的，即使是实践性很强的几何作图题，它强调的只是与公设一致的尺规作图．判断一个几何图形能否作出的惟一标准是看它能否经过逻辑推演归结为相关的公设以及由公设推演出来的一些基本作图，而不问人们在实际中能否运用其他工具和其他方法作出这个几何图形．

3．《几何原本》的历史意义

《几何原本》是古希腊数学的最高成就，它是一本最早的内容丰富的数学著作，它对数学思想的发展及人类文化的发展产生了深远的影响．

《几何原本》由于具有鲜明的直观性和严密的逻辑演绎方法相结合的特点，它被奉为数学教育的经典著作．人们正是从这本书里认识到数学是什么，证明是什么，公理演绎方法是如何的具有说服力、又是如何的优美，《几何原本》曾被翻译成各种文字的版本，在欧洲，直到18世纪，《几何原本》仍是学生的必修教材．在人类文化发展史中，被人们推崇了两千多年，而又影响这么长久的著作是少有的．

《几何原本》构造了数学史上第一个公理系统，它的产生不仅标志数学知识系统化的开始，而且开创了科学理论系统化的先河．公理化思想以及逻辑演绎推理思想所产生的深远影响，主要表现在以下几个方面：

(1)在公理化思想及逻辑演绎推理思想的形成和发展的同时，也推动了整个数学向前发展，从而产生了许多新的数学分支．例如，由于对《几何原本》公理系统中第五公设是否可由其他公设推出的研究创造了非欧几何；由于对公理系统协调性的研究，希尔伯特等数学家和逻辑学家创立了证明论或元数学；由于对形式系统与其相适应的模型之间关系的研究，使抽象代数与数理逻辑相结合产生了模型论等就是突出的例子．

(2)运用公理化思想和逻辑演绎推理思想去整理已有的数学知识，促使一系列数学分支的公理体系的创立，使各数学分支的表述更加科学化、系统化，逻辑基础更加严密．人们用公理化思想及逻辑演绎推理思想审视《几何原本》本身，发现它存在许多缺陷，为了弥补《几何原本》中的缺陷，许多数学家开展了重建欧氏几何的研究工作，在19世纪的最后30年中，数学

家们给出了种种几何公理体系，其中对概念阐述最精炼、最完备，其思想最接近欧几里得公理体系的是希尔伯特在1899年发表的名著《几何基础》。在这本书中不仅给出了欧氏几何的一个完备公理体系，而且提出和解决了公理化思想方法中若干理论问题，如公理系统的无矛盾性、独立性和完备性等。于是，《几何原本》中所有命题都能以希尔伯特公理体系为基础，经过逻辑演绎推导出来，这样，希尔伯特的《几何基础》不仅克服了《几何原本》中的所有缺陷，而且比《几何原本》上升到一个更高的层次，即由实质性的公理化思想方法上升到形式化公理思想方法(在实质性公理思想方法中，公理系统中的基本概念不是原始概念，而是给基本概念下了定义或确定了它的具体内容，且先于公理而被给定，公理只是表达这类特定对象的基本性质，而且必须是自明的；而在形式化公理思想方法中，基本概念是不加定义的，它的范围、内涵和特征是由公理体系确定的，并不是先于公理体系而确定，也就是形式化公理体系中的基本概念可以解释为各种不同的具体对象，只要它们符合公理规定的关系即可)。从此以后，许多数学分支如抽象代数、拓扑空间和集合论等都各自建立了自己的公理体系，有的还同时建立了几种不同的公理体系，如集合论中有著名的ZFC公理体系和NBG公理体系。特别值得指出的是柯尔莫哥洛夫于1933年建立了概率论的公理体系，使得像概率论这样与实际联系比较紧密的学科也能建立在牢固的逻辑基础之上，充分显示了公理化思想和逻辑演绎推理思想的威力。

(3)公理化思想及逻辑演绎推理思想对整个科学方法论的形成与发展起到了示范作用。自《几何原本》问世以来，不少人认为《几何原本》所使用的思想方法，不仅是建立几何学和整个数学体系的可靠的思想方法，而且也适用于建立所有的科学体

系．例如，在第一章中提到的17世纪的唯理主义学者斯宾诺莎曾经仿效《几何原本》，把人的思想、情感和欲望等当作几何中的点、线、面来研究，写出他的名著《伦理学》．现代科学（尤其是自然科学）更是尽量采用形式化公理思想方法作为研究和表达手段．例如，拉格朗日在1788年发表的《分析力学》一书中，从虚功原理出发进行纯数学演绎（不借助任何实验和图形）得出了分析力学中的主要结论，奠定了现代公理化理论力学的基础．又如现代科学巨星爱因斯坦是一位精通几何学并且应用公理化思想及演绎推理，开创自己研究工作的物理学家．他多次提出在物理学的研究中也应当在逻辑上从少数几个所谓公理的基本假定出发．他在1905年创建的狭义相对论中就运用了这种思想方法，把狭义相对论的整个理论建立在两条公理上：相对性原理和光速不变原理．爱因斯坦的这一理论成果对推动现代物理及科学技术向前发展起到了巨大的作用．

二、《九章算术》及其影响

1.《九章算术》形成的历史背景

《九章算术》是我国古代一部数学专著，它不是一个人的作品，也不是一个时代成书的，而是经过许多数学家的修订和增补才逐渐成书的．

从春秋战国到秦汉时期，随着铁器的使用、矿产的开掘、大型水利工程的兴修等，使我国手工业、农业及各种技术有了很大的发展，数学也获得相应的发展．事实上，生产与生活都需要各式各样的数学知识，或提出新的数学问题．例如，西汉时期由于农业发展，生产的粮食用不完，需要修建大量的粮仓贮存粮食，从而需要计算各种立体的体积．各种粮食的交换，又规定了比例关系，汉武帝时又实行“均输法”，从数学角度看

就是由国家规定合理运输赋谷的方案，汉代规模浩大的水利工程则需要更复杂、更精确的计算．所有这些，使得这个时期的数学得到充分的发展，内容日益丰富，并且具有算法化和实用性的倾向．

另一方面，在春秋战国时期，儒、法、名、阴阳、墨等"诸子百家"形成了不同的流派，他们在政治上和学术上都各自提出自己的主张，百家争鸣，学术气氛异常活跃，其中名家和墨家都会运用形式逻辑，特别是墨家更为突出．但在墨家之后、刘徽之前的大约六七百年之间，形式逻辑在中国不但没有向前发展，而且处在衰落时期，在这种特有的文化背景下形成的中国古代数学体系自然具有自己鲜明的特征——非逻辑化倾向．

将数学用于解决实际问题及非逻辑化的学术思潮对中国古代数学的发展产生了极大的影响，《九章算术》正是在这种背景下成书的，它与古希腊数学的逻辑化、几何化倾向相反，具有算法化、实用性及非逻辑化倾向．

2.《九章算术》的主要内容和思想特点

《九章算术》是我国古代数学的一部经典著作．它初步形成了以计算为中心、以解决实际问题为目的、形数结合的归纳体系，有着丰富的数学知识内容和独特的思想特点．

《九章算术》中共列出246个实际问题，199条术文，共分为九章．第一章名为"方田"，列题38个，立术21条，主要是讲平面几何图形的计算方法．第二章名为"粟米"，列题46个，立术33条，主要是讲述以谷物交换为中心的比例算法．第三章名为"衰分"，列题20个，立术22条，主要是讲以分配为中心的比例分配．第四章名为"少广"，列题24个，立术16条，讲述已知矩形(含正方形)面积求一边之长和已知立方体体积求

其棱长的开方法则，以及已知球的体积求其直径的算法．第五章名为"商功"，列题28个，立术24条，讲述以各种形状的立体的体积计算公式．第六章名为"均输"，列题28个，立术28条，讲述以赋税计算和其他应用题为中心的比较复杂的比例计算方法．第七章名为"盈不足"，列题20个，立术17条，讲述以盈亏问题为中心的一种双假设算法．第八章名为"方程"，列题18个，立术19条，讲述以多元问题为中心的线性方程组的算法(相当于今天的矩阵初等变换算法)．第九章名为"勾股"，列题24个，立术19条，讲述以测量问题为中心的直角三角形的三边互求，以及容圆、容方的解法，还讲到相似勾股形、勾股数和一元二次方程的解法．

《九章算术》中的数学思想主要表现在下列几个方面：

(1)经世致用思想

《九章算术》是按当时社会实践所需要解决的问题来分类的，每一类中设置若干实际问题．每个问题都提供有关的算法和问题的答案．或是从大量实际问题的计算中提炼出来的数学模型．由此可以看出，《九章算术》中建立了一个按照实际应用问题性质归纳分类的开放性的理论体系，这种理论体系与《几何原本》中建立的封闭的逻辑演绎体系不同，它对整个社会生产和生活实际是开放的，其内容和体系都与当时的社会生产和生活实际紧密相关．在《九章算术》成书之后的二千多年时间内，它成为了中国古典数学的经典著作和范本．《九章算术》对于现代数学思想发展具有重大的影响，事实上，现代应用数学正是按照应用方向或主要采用的数学模型来分类的．

(2)算法化思想

《九章算术》中的思想方法集中表现在它所给出的几十个数学公式和法则上，它们可归纳为四类算法和两类求积公式．

四类算法是分数算法("方田"章)、一般比率算法("粟米"、"衰分"、"均输"三章)、开方算法("少广"章)以及组合比率算法("盈不足"、"方程"、"勾股"三章)．两类求积公式是面积公式("方田"章)和体积公式("商功"章)．不少算法给出了程序化的计算步骤，因而都可以编成程序在现代电子计算机上实施，充分体现了算法化(机械化)的思想特征．即把满足一定条件的较复杂的数学计算，分解为简单的、刻板的、重复的机械动作，达到以数目较多的一些简单的量的工作去实现比较复杂的质的目的．

(3)形数结合思想

《九章算术》中包含有丰富的几何内容，分别属于"方田"、"少广"、"商功"、"勾股"等章．在这些章节中对图形的研究表现为数量的计算，它以长度、面积和体积等度量为主要计算对象，充分体现了以"计算为中心"的思想特征．例如刘徽注的《九章算术》中把相似勾股形的概念作为基本概念，把相似勾股形的性质作为基本性质，通过相似勾股形的性质把一些几何问题转化为代数问题，并与长度、面积和体积联系在一起，推导出一系列计算公式，形成了求积理论中的相比法和勾股术中的勾股比率论．另一方面，在《九章算术》中，几何的原理和方法被成功地应用于代数．例如《九章算术》中的开平方术、开立方术和二次方程的数值解法都是来源于几何，此外，刘徽还用出入相补的几何方法得出了求整勾股弦的一般公式．由此可以看出几何方法与代数方法的互相渗透、数与形的巧妙结合成为了《九章算术》的又一个显著的思想特征．

(4)模型化思想

《九章算术》中每一章都是从大量实际问题中选择具有典型性的现实原型作为所讨论的问题，然后通过"术"(即算法)用形

式化语言对问题中各元素的关系、问题的本质及解决问题的一般过程进行描述，概括地表达出一定的数学结构，也就是将实际原型转化成为数学模型．当然有的则先给出数学模型，再举出这些数学模型可以应用的原型．通过对数学模型的研究得出算法来．而算法运用于数学模型表述出来的一类问题，以原型中的处理方法为范例，人们就可以应用所得出的算法去解决实际中的问题．可见模型化思想是与算法化思想紧密相关的．针对建立起来的数学模型去研究与之相适应的算法也正是近代计算数学理论所要解决的主要问题．

(5)辩证思想

中国古代哲学的辩证思想极为丰富，而中国古代数学植根于社会生活实践之中，必然受到中国古代哲学的深刻影响，这种影响在《九章算术》中也得到了充分的反映．刘徽在《九章算术》中多次运用了极限观念及其理论，正确地解决了一些较难的数学问题，他从圆内接 6 边形算起，令边数一倍一倍地增加，逐个算出 6 边形、12 边形、24 边形…的面积，逐步逼近圆面积，从而求出圆面积．他算到圆内接正 192 边形，得到 π 的近似值是 3.141024，且他还证明了圆田术："半圆半径相乘得积步"，也就是圆面积等于半圆周长与半径之积．此外，刘徽的这种极限思想还可在他关于开方不得整数的平方根的近似值的算法讨论中看到．在数学中应用极限思想是刘徽所注的《九章算术》中的一项重大成就．只有具有中国古代的那种包容一切、对立的两极在运动中发生转化的辩证思想，才可能理解正多边形与圆的对立统一．刘徽在《九章算术》注中虽然没有形成近代数学极限理论中有关数列、数列和、单调数列、有界数列以及收敛数列等确切概念，但是，刘徽所使用的数学思想方法与近代数学极限理论中的思想方法是基本一致的．此外，

《九章算术》中提出了正负数的概念以及关于正负数的运算法则．从《九章算术》的形成的时代算起，中国人使用“正”、“负”数术语至今天至少有1500年以上的历史了．中国之所以能先于西方十几个世纪引入负数并承认了负数，关键在于中国古代数学的辩证思想．

§2.2　非欧几何的诞生

19世纪，随着数学的发展，产生了一门新的几何学——非欧几何学．它突破了两千多年来传统的欧氏几何的束缚，给传统的数学思想以猛烈的冲击，不仅推动了几何学向前发展，而且对现代数学、现代自然科学和数学哲学产生了深远的影响．

一、非欧几何思想的产生

非欧几何思想的产生可溯源于《几何原本》中的第五公设．由于第五公设(即平行公理)不仅叙述文字啰嗦，而且不符合亚里士多德关于公理“自明性”的要求，所以自欧几里得以来，人们都试图用其他公设和定理去证明它．意大利的萨开里是第一个试图用反证法证明第五公设且影响较大的数学家．他的做法是：如图2－1，假设四边形$ABCD$中，$\angle A=\angle B=$直角，$AD=BC$．容易证明$\angle C=\angle D$．为了证明$\angle C=\angle D=$直角(从而可证明“第五公设”)，只要否定下列两个假设：(1)$\angle C=\angle D=$钝角；(2)$\angle C=\angle D=$锐角．萨开里从(1)出发很快引出了矛盾，于是否定了(1)，但是他从(2)出发却推出了很多结论．其中第38个推论是：在平面内存在两条直线l_1和l_2，它们在一个方向无限地互相接近，而在其相反方

向无限地分开，这样 l_1 和 l_2 将在无穷远点处有公共的垂线。他说这显然是不可能的，是"矛盾"，其实这只是与人们的视觉经验相矛盾，并没有与已有的其他公设和由这些公设推出的定理产生矛盾。因此，萨开里自认为证明了第五公设，其实并没有证明。

瑞士的几何学家兰伯特在 1766 年写成了《平行线论》一书，他沿着萨开里的思路走下去，并没有因为从(2)推出的结论不符合人们的直观经验就轻易否定它们，而是对平行公理的可证明性提出了怀疑，他大胆猜测：从(2)推出的一系列结论可在欧几里得空间的虚球面上实现。这是一种数学思想上的重要突破，没有这种思想上的突破，非欧几何就不可能诞生。比兰伯特更进一步的是德国数学家施卫卡特和托里努斯，他们不但正确地猜测到平行公理的不可证明性和新几何的存在(施卫卡特称这种新几何为星空几何，因为他猜测这种几何可在星际空间成立)，而且还为新几何进行了某些探索性工作，相当于非欧几何的一些粗略观念。

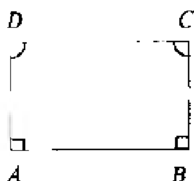


图 2－1

非欧几何的最终创立应归功于三位数学家，他们是高斯、亚·鲍耶和罗巴切夫斯基。从时间上来看高斯最早，他在 1816 年左右就已经提出了非欧几何的基本思想，确信存在一种与欧氏几何不同的几何学，尽管他已认识到非欧几何的深远意义，但他害怕新的几何理论不会被人理解，而会被人嘲笑，因此，终生没有向世人公开宣布自己的研究成果。不仅如此，对于别人将非欧几何的研究成果寄给他，请他评审时，他也没有公开表示支持，高斯关于非欧几何的思想是在他死后从他给朋友的信和遗稿中才被发现的。

亚·鲍耶是高斯的朋友法·鲍耶的儿子，他写成的关于非欧几何的论文是作为其父著作的附录发表的，附录的题目是"附录：叙述空间绝对真理性的学说，这个真实性与欧几里得的第十一公理的真伪无关…"，当法·鲍耶把儿子写的附录寄给高斯，请他评论时，高斯回信说："…称赞他等于称赞我自己，因为这里研究的一切内容，你儿子所采用的方法和他所达到的一切结果，几乎和我的一部分在30—35年前已开始的个人沉思相符合，我真是被这些吓坏了…"信中并表示他不便公开赞许，虽然，亚·鲍耶在非欧几何发现面前并无动摇和软弱表现，但他在上述附录中提出的概念和命题太过于简练，以致没有引起学术界的注意。

第五公设问题的彻底解决者是罗巴切夫斯基，他在1826年2月11日喀山大学物理数学系会议上，宣读了报告"关于几何原理的讨论"，并于2月23日又宣读了报告"平行线理论和几何学原理概论及证明"，之后，他还公开发表了"几何学原理"(1829年)，"平行线理论的几何研究"(1840年)等一系列有关非欧几何的论文，这样，罗巴切夫斯基开辟了几何学的一个新领域，创建了非欧几何学，从而从理论上解决了第五公设不可证问题，罗巴切夫斯基创立的新几何动摇了旧的传统观念，因而引起教堂的反对，总主教宣布他的学说是邪说，有人用匿名信在杂志上漫骂、嘲笑、侮辱他，甚至宣布他是疯子，这一切正如高斯所预料的，也是高斯害怕的，高斯仅在私人通信里说到自己对罗巴切夫斯基理论的钦佩，而不敢公开支持他，但是罗巴切夫斯基却不屈服，坚持真理，一个人英勇奋斗到底。

从上述情况来看，创立非欧几何的三位数学家，在思想上所经过的历程大致相同，又几乎是同时在三个不同的国家得到类似的结论，这表明一种新的数学思想在经过长期积累后必然

会出现．即使不是高斯、鲍耶和罗巴切夫斯基，也会有别的数学家创立非欧几何．其次，非欧几何的创立还说明创立一种新的数学思想，必须敢于同顽固的传统思想和传统观念决裂．高斯正是由于缺乏罗巴切夫斯基的这种精神而不敢公开自己的研究成果，致使非欧几何的诞生至少推迟了10年．

1854年，德国青年数学家黎曼在哥廷根作了"关于作为几何学基础的假设"的著名演讲(1868年正式发表)，提出了更为一般的几何空间——广义黎曼空间，作为特例便得到普通欧氏空间、罗巴切夫斯基空间(双曲空间)和狭义黎曼空间(椭圆空间)，黎曼在报告中所阐述的几何思想极为深刻，据说当时听报告的人中除了年迈的高斯外没有一个人听得懂．黎曼之后，意大利数学家贝尔特拉米于1866年发表了"非欧几何解释的尝试"一文，证明了非欧几何可以在欧几里得空间中伪球面(负常数高斯曲率曲面)上实现．接着，1870年德国数学家克莱因也给出了一个解释，之后法国数学家庞加莱又给出了一个解释．通过他们建立的模型可以将非欧几何中的命题"翻译"成欧氏几何中的相应命题．既然，人们承认欧氏几何中没有矛盾，自然也就承认了非欧几何中也没有矛盾．至此，非欧几何才得到学术界的承认和赞扬，一场长达两千年的关于第五公设是否可证明的争论才得到圆满地解决．

二、非欧几何创立的历史意义

非欧几何的创立打破了两千多年以来欧氏几何一统天下的局面，从根本上革新和拓宽了人们对几何空间的认识，给数学开辟了研究各种空间的新领域．1872年克莱因提出了著名的爱尔朗根纲领，用变换群的观点对各种几何学(包括欧氏几何、罗巴切夫斯基几何、黎曼几何等)进行了分类，对几何学的发

展产生了重大的影响．之后黎曼几何又有了新的发展，产生了更一般的以曲线长度积分为基础的芬斯勒空间，以超曲面的面积分为基础的嘉当空间，以二阶微分方程为基础的道路空间，…这些统称为一般空间．比黎曼空间的概念更广．因此，可以说，非欧几何的诞生是数学由近代数学时期向现代数学时期转折的重要标志之一，也是数学从以直观为基础的时代进入以理性为基础的时代的重要标志之一．

1．非欧几何所蕴含的数学思想是非常深刻的，它的产生表明数学的逻辑结构对现实直观具有相对独立性，也表明数学的逻辑推理可独立于物质世界进行，从而把人们的思想从"数学结论必须符合感性直观"这一传统观念的束缚下解放出来．它的成功还表明，人们的数学思想不能固守在某个模式之中，一旦从正确的前提出发，经过正确的逻辑推理得出了新的数学思想方法和新的结论，就要敢于突破旧的思想的束缚，坚持真理，才不致于像高斯那样发现了真理而不敢于坚持的遗憾．

2．非欧几何的创立促使了公理化思想方法的完善和发展．公理化思想方法经历了三个阶段：公理化思想方法的产生阶段、完善阶段以及形式化阶段．第一阶段包括亚里斯多得的公理化思想以及欧几里得的《几何原本》．由于《几何原本》中所论述的是一些特定的几何对象(点、线、面等)，这些对象具有明显的直观背景，因此，《几何原本》中的公理体系可称为"实体化公理"．第二阶段包括对欧几里得公理体系的各种思考(其中以对第五公设是否可证明的思考最引人注目)，这个时期经历了两千年，直到非欧几何的创立．第三阶段则以希尔伯特的名著《几何基础》的问世为标志．希尔伯特公理体系中称为点、线、面的对象，并不具有任何具体的意义，这些名称甚至可换为任何其他术语，这样就完全摆脱了空间观念的直观部分．因

此，《几何基础》中的公理体系可称为“形式化公理”．在上述三个阶段中非欧几何的创立具有承上启下的作用，它对公理化思想方法的完善和进一步发展产生了深刻的影响，具体表现为下列几个方面：

（1）非欧几何表明，第五公设是不能由其他几个公设或公理来证明的，它是独立的命题．这使人们认识到一个完备的公理体系中每一个公理都必须是独立的，我们可将其中一个独立的公理用其他独立的公理来代替而形成一个新的公理体系，这是现代数学中一种重要的思想方法．

（2）非欧几何中一系列结论与人们的直观经验相矛盾，人们自然要问：将欧氏几何中第五公设用罗巴切夫斯基或黎曼的平行公理代替后，与其他公理是否矛盾呢？虽然贝尔特拉米、克莱因和庞加莱用数学模型方法证明了非欧几何相对于欧氏几何是不矛盾的，但是这不能使人们放弃对欧氏几何公理体系本身的相容性的探讨．由于有了解析几何，这等于在实数系统中构造了一个欧氏几何的模型．于是，欧氏几何的相容性就归结为实数系统的相容性，而实数又可用有理数的分割来定义，于是又归结为自然数系统的相容性，而自然数系统或算术系统的相容性又归结为集合论的相容性．算术公理的相容性问题是希尔伯特在1900年世界数学家大会上提出的著名的23个问题中的第二个问题，这个问题虽然取得了一些进展，但至今尚未完全解决，因此集合论的相容性问题至今也还是一个谜．当然，如果假定算术系统是相容的，那么欧氏几何的相容性就可以得到证明．可见，对非欧几何相容性的研究促进了公理化思想方法的深入研究，也导致了人们对数学基础的研究．此外，证明非欧几何相对相容性的数学模型方法也成为了现代的一种重要的数学思想方法．

(3) 非欧几何的创立，是对欧氏几何公理体系的反思而引起的，这使后来的数学家注意对几何基础乃至整个数学基础的研究，除了上述关于公理体系的独立性和相容性的探讨以外，人们还致力于公理体系的完备性研究．所谓完备性指的是这个公理体系中的公理已构成最广义的集合，不允许再添入其他新的公理(大家早已发现欧几里得原来的公理体系并不完备)． 1871－1872年间德国数学家康托和戴德金不约而同地提出了连续性公理，1882 年德国数学家巴许又提出了顺序公理，巴许还从理论上提出了形式化公理方法的思想．正是在许多数学家的研究工作基础上，希尔伯特提出了著名的希尔伯特公理体系，这个体系不仅满足独立性和相对相容性(即相对算术体系是相容的)等要求，而且希尔伯特利用模型论中的重要概念——范畴性证明了他提出的公理体系是完备的．事实上，希尔伯特根据他提出的公理导出了欧氏几何的若干基本定理，后人又继续演绎推导，直到证得欧氏《几何原本》的所有定理，充分显示了希尔伯特提出的几何公理体系已经排除了《几何原本》的不足之处．此外，希尔伯特已经把公理化思想方法发展到了形式化的阶段，并且这种形式公理化思想方法迅速从几何领域扩展到算术、数理逻辑、集合论、概率论等领域，由此形成的公理化思想方法，已经成为现代数学的重要思想方法之一．

3．非欧几何的创立，改变了欧氏几何是描述物质空间惟一真理的看法，这对于 20 世纪初物理学中关于空间和时间的物理观念产生了重大的影响．爱因斯坦为了建立广义相对论的精确表达方法，不得不花费 7 年的时间来钻研非欧几何学，至 1915 年，他终于阐明了广义相对论中关于"引力场"的几何学理论．按照相对论的观点，宇宙结构的几何学不是欧氏几何学而恰恰是接近非欧几何学．

4．非欧几何的内容逐步渗透到其他数学分支．例如，黎曼几何不仅是微分几何的基础，而且运用到微分方程、变分法和复变函数等方面．法国数学家庞加莱利用复半面作出的罗巴切夫斯基几何(双曲几何)模型证明了自导函数的基本区域是一些彼此全等的多边形，对建立自导函数的基本理论起了重大的作用．

5．非欧几何的创立，不仅引起了一些重要数学分支的产生，促进了数学的发展，而且使数学家加深了对数学本质的理解，使数学哲学的研究进入了一个崭新的历史阶段．在如何看待数学的真理性这一问题上，由于非欧几何的诞生，使许多数学家和哲学家提出了各种不同的观点，长期争论，逐渐形成了各种不同的数学真理论，如经验真理论、分析真理论及条件真理论等等．此外，数学基础问题、数学本体论问题以及数学中悖论问题等一系列数学哲学问题都与非欧几何有着不可分割的联系．

总之，正如希尔伯特所说的："19 世纪最有启发性、最重要的数学成就是非欧几何的发现．"

§2.3　三次数学危机

一、$\sqrt{2}$的发现与第一次数学危机

1．第一次数学危机的产生

公元前 5 世纪，古希腊的毕达哥拉斯创立了一个崇拜阿波罗神和从事数学研究的秘密社团，历史上称为毕达哥拉斯学派． 他们认为数是万物的本原，数产生万物，数的规律统治万物，用毕达哥拉斯的话来讲，就是"万物皆数"．毕达哥拉

斯和他的门徒并通过数的抽象研究去论证他们的"和谐的宇宙系统"的世界观。

毕达哥拉斯学派对几何学的贡献最大。他们发现了所谓毕达哥拉斯定理，即直角三角形两直角边的平方和等于斜边的平方。我国称此定理为勾股定理或商高定理。商高比毕达哥拉斯要早600多年知道这一结论。毕达哥拉斯(或他的门徒)发现，当直角三角形两直角边相等时，斜边与直角边之比是不可公度的，即斜边与直角边的比值$\sqrt{2}$不再是数(指自然数或分数)，据说他们还用归谬法证明了这一结论。本来这一发现是毕达哥拉斯学派的一大成就，因为这一事实不能从观察和经验得出，只能通过逻辑推理和抽象的思考来证明，这说明人们的抽象思维能力有较大的提高。但这一发现引起了毕达哥拉斯学派的惶恐不安，因为当时人们的认识刚刚从自然数扩充到有理数，凭经验他们认为"一切量皆可用自然数或分数来表示"。可是居然发现等腰直角三角形中当直角边等于1时，斜边这个实实在在的量却不能用数表示，岂不与他们所奉行的"万物皆数"的信条相矛盾吗？这一发现动摇了毕达哥拉斯学派的哲学偏见，使他们陷入极度不安的深渊之中，因而产生了数学基础的第一次危机。

2．第一次数学危机的影响

数学史上的第一次危机对数学乃至其他科学的发展都产生了深远的影响。希腊人在解决这次危机中取得了一系列重要成果。

首先，古希腊人为了继续维护自己"万物皆数"的信仰，拒绝承认$\sqrt{2}$是数，而称它为几何量，把数和量加以割裂，认为数是离散的，几何量是连续的，是不可公度的。于是，希腊学者使用了包括不可公度在内的比例论来克服这个困难。这项成

果被欧几里得收集在《几何原本》第五卷和第十卷中．比例论似乎能使希腊人发现无理数，并由此建立研究连续变化的算术理论——实数理论．然而，由于他们受哲学上的偏见所禁锢，始终没有做到这一点．但是他们为了克服这次危机而发展的比例论却是欧氏几何中的最大成就之一．

其次，由于无理数的发现导致了对连续量的研究，而当时的算术理论中没有连续量，这就必然以几何量的连续性这种直观概念为依据，也就产生了用于处理连续量的穷竭法，一般称它为阿基米德预备定理．虽然穷竭法在当时还没有充分推广，但它可以看成是近代积分理论的先驱．

此外，这次危机使古希腊人在数学研究方向和数学思想方法上产生了巨大的变化．由于几何量不能完全由自然数及其比表示，而任何数却可由几何量表示，故自然数在人们心目中的地位动摇了．希腊人开始转向偏爱几何学．又由于直觉和经验不一定靠得住，推理和证明才靠得住，从此希腊人开始重视几何学的演绎推理，导致欧几里得《几何原本》的出现，从而建立了几何学的公理体系；同时由于注重推理，产生了亚里斯多德的名著《工具篇》，创立了逻辑学的公理体系．这两个历史上最早的公理体系是几乎同时诞生的，可以说这是数学思想及科学思想上的一次重大革命．它们被视为金科玉律与推理楷模，享受了两千多年的盛誉，直到19世纪末才得到本质上的改造和完善．

二、无穷小与第二次数学危机

1．第二次数学危机的产生

17世纪下半叶微积分的出现不仅为近代数学的发展提供了基础，而且为近代科学技术的发展提供了十分重要的数学方

法．此外，微积分也引入了一些传统观念无法理解的概念和方法．由于引入的一些概念带有模糊性，无法从逻辑上作出一致的解释，从而引起长期的、尖锐的争论，产生了数学基础的第二次危机．

下面我们以牛顿的微积分思想为例来说明这次危机是如何产生的．牛顿在1669年写了一本小册子《运用无穷多项式方程的分析学》(1711年正式出版)，其中有下列例题：设有一条曲线，其下面积为$z=ax^m$(m为有理数)，当横坐标x获得一个无穷小增量"0"(牛顿称为"瞬"(monent))，相应的面积z有一个无限小的增量"$0y$"，于是

$$\begin{aligned} z+0y &= a(x+0)^m \\ &= a\left(x^m + mx^{m-1}\cdot 0 + \frac{m(m-1)}{2}x^{m-2}\cdot 0^2 + \cdots\right). \end{aligned} \qquad ①$$

考虑到$z=ax^m$，抵消后两边除以0，得

$$y = max^{m-1} + a\left(\frac{m(m-1)}{2}x^{m-2}\cdot 0 + \cdots\right). \qquad ②$$

略去含0的项得

$$y = max^{m-1}. \qquad ③$$

这就是对应于面积Z的曲线的纵坐标y的表达式，结果表明：若面积由$z=ax^m$给出，那么构成这个面积的曲线是$y=max^{m-1}$；反之，如果曲线是$y=max^{m-1}$，那么，它下面的面积是$z=ax^m$．这实质上就是微积分学的基本定理．

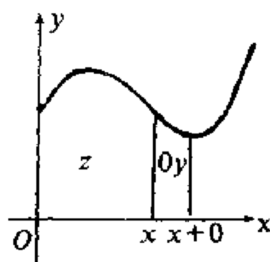


图2-2

在牛顿的计算过程中，“0”到底是什么？他没有作解释，从他的证明过程来看，他先把“0”看成非零的常数(尽管它很小)，否则“横坐标变为 $x+0$”，两边除以“0”没有意义，无法从①式推出②式．最后，他舍去含有“0”的因子的各项，又把“0”视为零，否则无法从②得出正确的结论③，在这里牛顿遇到了逻辑上的严重困难．

显然，微积分的发明得到了广大数学家和物理学家的热烈欢迎，大家充分运用它解决了许多过去无法解决的科技难题，但也由于它逻辑上的漏洞，招来了一批数学家和哲学家的抨击甚至讽刺和指责．对微积分方法攻击最为厉害的是英国主观唯心主义哲学家大主教贝克莱，他在1734年出版的《分析学家，或者向一个不信正教的数学家的进言》一书中认为牛顿的无穷小量“0”开始不是零，然后又让它等于零，违背了矛盾律．他甚至讽刺说既然无穷小量不是零又不是非零的量，那它一定是量的鬼魂了．相信量的鬼魂的数学家们，又有什么理由不相信上帝的存在呢？贝克莱的攻击虽然出于为宗教和神学辩护，但他对微积分的非难却切中要害，从而正确指出了牛顿和其他数学家们对自己的推理方法没有给出严格的逻辑解释，数学家和物理学家相信它，仅仅只是因为它使用起来十分有效，并且结果总是正确的．这样，整个18世纪，数学家和哲学家们就围绕微积分的无穷小悖论展开了一场大论战，曾给数学界带来了诸如乱用符号、乱用无穷级数的混乱局面，被史学家称为数学基础的第二次危机．

2．第二次数学危机的影响

数学基础的第二次危机的影响是深远的，正是在解决危机的过程中，数学获得了巨大的发展．

首先，第二次危机表明微积分的基本理论还有待发展和完

善．在给微积分建立严格的理论的努力中，许多数学家做了大量的工作，其中以柯西、魏尔斯特拉斯等人的工作最为突出．通过他们的努力，才使分析学有了严格的理论基础．

其次，为了解决第二次数学危机，促使数学家进一步深入研究微积分的基础——实数理论．直到19世纪70年代，魏尔斯特拉斯、康托、戴德金等数学家分别独立地建立了各种不同的实数理论．而极限理论又是以实数理论为基础的，从而使微积分建立在严格的实数理论的基础上，同时也导致集合论的诞生．此外，为了解决这次危机，传统的形式逻辑变化为数理逻辑，使数学推理有了更广泛的逻辑基础．

最后应当指出的是，极限理论的建立虽然解决了无穷小量悖论，但极限理论的相容性仍未得到证明．因为极限理论的无矛盾性归结为实数理论的无矛盾性，而实数理论的无矛盾性又归结为集合论的无矛盾性，但集合论的无矛盾性至今仍未彻底解决．因此，我们说无穷小悖论已经得到解决，指的是在一定条件下得到相对意义下的解决．

三、罗素悖论与第三次数学危机

1．第三次数学危机的产生

从前面的论述可以看到，欧氏几何公理体系的相容性、极限理论的相容性以及实数理论的相容性都归结为集合论的相容性．德国数学家康托于1874－1897年间发表了一系列关于无穷集合的论文．1872年的论文除了用相当大的篇幅叙述以基本序列定义无理数的实数理论并提出"戴德金—康托公理"外，还定义了"导集"概念，并以导集的性质为准则对无穷集进行了分类，使无穷点集成为了数学研究的对象．1874年，他在"关于一切代数实数的一个性质"一文中提出了"可数集"概念，并

以一一对应为准则对无穷点集进行了分类，证明了一切代数数是可数的；任何有限线段上的实数是不可数的；超越数是不可数的；一切无穷集并非都是不可数的：无穷集同有穷集一样也有数量(基数)的区别等．这篇论文标志着朴素集合论的诞生．随后，他又独立定义出超穷“基数”和“序数”等概念，经过多年的努力，他终于建立了具有划时代意义的抽象集合论．康托的集合论成果使数学界喜气洋洋、一片欢乐．数学家庞加莱在1900年召开的第二届国际数学家会议上所作的报告中得意地宣称：“数学的严格性，看来直到今天才可以说是实现了”．可是不到两年，英国著名的哲学家、数学家和社会改革家罗素提出了著名的罗素悖论，指出集合论中存在矛盾，并不相容．罗素悖论的内容是：设 T 是一切不以自身为元素的集合组成的集合，要问 T 包含还是不包含自身呢？(1) 如果 T 不包含自身，即 $T \in T$，那么依 T 的定义又应有 $T \in T$．这就得到矛盾；(2) 如果 T 包含 T，即 $T \in T$，那么按 T 的定义又应该 $T \in T$．同样得到矛盾．罗素悖论仅用到“集合”、“元素”、“属于”这些最基本的概念，而且引进集合的原则符合通常的方法——用描述集合元素性质的方法来定义集合．可见罗素悖论的发现使数学中最基本的概念和最重要的推理方法受到威胁，从而使整个数学基础发生了动摇，整个数学界为之大惊！正因为悖论的发现在数学界引起了如此大的震动，所以数学史上称它为数学基础的第三次危机．

2．第三次数学危机的影响

由于悖论引起的第三次数学危机推动数学沿着两个方向发展：一个是单纯从集合论自身考虑，怎样才能从集合论中排除悖论；另一个则是从整体上考虑整个数学科学的对象、方法和理论结构．

第一个方向的直接结果是导致集合论的公理化。本世纪以来，许多数学家为了从集合论中消除悖论，提出了种种方案，其中最主要的是ZFC公理体系和NBG公理体系。

ZFC公理体系是由策墨罗于1908年提出，后经弗朗克尔和斯科朗多次修改和补充而成的。策墨罗认为集合论之所以出现悖论，是因为使用了太大的集合(如包含所有集合的集合)。因此，为了避免悖论必须对集合加以限制。为此，必须对康托等人提出的"集合概括原则"(即每个条件可以确定一个集合，亦即每个概念的外延可以确定一个集合)加以改造。策墨罗提出用"有限抽象原则"代替"概括原则"(也称无限抽象原则)，这个有限抽象原则是：如果给定了一个集合，又给了一个条件，那么这个集合中所有满足那个条件的元素可以构成一个集合。这个原则实质上说，事先应该给概念的外延划定范围，由此消除了康托定义造成的任意构造集合的现象，从而避免了已发现的一些悖论。

NBG公理体系是由冯·诺伊曼首先提出，后经贝尔耐斯和哥德尔改进而成的。冯·诺伊曼认为集合论中产生悖论的原因并不是由于使用了太大的集合，而是由于允许集合可以作为自身的元素(即 $S \in S$)产生的。所以只要规定不允许集合可作为自身的元素就行了。

ZFC公理体系和NBG公理体系都消除了已出现的悖论，并且至今在这两个公理体系中还没有发现悖论。但这种无矛盾性还没有得到严格的理论证明，而且根据哥德尔不完全性定理，ZFC和NBG系统本身都不可能证明自己是无矛盾的，即它们的无矛盾性都只能借助外系统来证明。尽管如此，公理化的集合论的建立对纯粹数学、特别是数学基础的研究，奠定了巩固的逻辑基础。

第二个方向的直接结果是导致对数学基础的研究．在怎样彻底消去集合论中悖论的途径上，许多数学家已经注意到，仅用公理化集合论的思想方法并不能完全解决问题，还有许多问题需要深入研究．例如，集合论的矛盾是从哪里来的？如何理解“数学的存在”？数学的基础是什么？等等．对这些问题的不同回答，形成了数学中的不同学派，引起了20世纪初关于数学基础的一场大论战．其中最主要的三个学派是逻辑主义、直觉主义和形式主义学派(参看第一章)．

由于三大学派之间的争论暴露出的是整个数学性质的问题，因而对数学的发展具有十分积极的意义，并对现代数学思想产生了深远的影响，使人们对数学本身产生了一些新的认识：

(1) 解决数学问题的关键是正确认识和处理有限和无限．

为了消除悖论，三大学派都在寻找产生悖论的根源，思想都集中在处理“无限”上．数学中“无限”这个概念，自古以来一直是一个争论的焦点．古希腊人用穷举法回避它，后人又用极限、非标准分析的理论来处理它．正如克莱因在《古今数学思想》(第四册)一书中所说的那样：“1930年以后的全部发展还留下来两个未解决的大问题：去证明不加限制的经典分析与集合论的相容性，以及在严格直观的根基上去建立数学，或者去确定这种途径的限度．在这两个问题中困难的根源都在于无穷集合和无限程序中所用到的无限．”

(2) 数学的严格性不是绝对的

为了排除数学中的悖论，使数学能奠定在更加牢固的基础上，许多数学家都认识到相容性(即无矛盾性)是一个非常重要的问题．希尔伯特用有穷推理方法来研究形式系统的相容性、完备性及判定性等问题，并试图通过研究形式系统的相容性来

说明数学理论的相容性．但是哥德尔证明了："如果系统只限于使用能在数论系统中形式表示的概念和方法，那么，包括通常逻辑和数论的系统的相容性是不可能建立的"．因此，希尔伯特试图用证明论中所许可的狭义逻辑来建立数论系统的相容性是不可能的．后来，根岑放松了希尔伯特证明论中所允许的证明的方法的限制，于1936年使用"超限归纳法"证明了数论系统的无矛盾性．可见，相容性是有条件的，它是绝对和相对的对立统一．罗素发展了"类型论"来消除悖论，并试图在这个基础上构造出一个逻辑公理体系，从逻辑学推出全部数学，从而得出数学的无矛盾性的结论．又如直觉主义学派仅用构造的方法产生数学，摈弃了古典数学中很大的一部分，试图使留下的部分是不矛盾的．然而，哥德尔严格证明了当一个形式系统复杂到一定程度时，如果它是无矛盾的，那么就存在不可判定的真命题，即它一定是不完备的．可见形式系统的无矛盾性得到证明，但不能反映全部数学理论的无矛盾性．并且哥德尔还证明了，如果古典数论有互相矛盾的定理，那么直觉主义数论的解释也是互相矛盾的，等等．总之，无矛盾性和无矛盾性的证明都是有条件的，数学的严格性不是绝对的，正如克莱因在《古今数学思想》(第四册）一书中所说的那样，"严格性在希腊时代就变成了一个目标，虽然到19世纪以前受到冲击时仍受到尊重，它似乎要达到了，但是，过分追求严密性将引入绝境而失去它的真正意义．"

三大学派的失败表明了，单靠公理化思想方法、逻辑推理思想方法或构造性思想方法来推动数学发展是不够的，因为数学面临的矛盾不仅有数学内部的矛盾，而且有数学与现实的矛盾．因为现实世界本身是对立统一的，数学只能在绝对和相对的对立统一中建立其无矛盾性，试图要为数学奠定一个绝对可

靠的基础是不可能的，"数学的真理性并不存在于形式演绎系统的严格证明里，而归根结底要通过与物质世界相联系的实践过程去验证"(徐利治，《数学方法论选讲》，华中工学院出版社，1983年，第175页)．

§2.4 近代和现代数学思想的几次重大变革

一、变量思想进入数学——解析几何与微积分的产生及其意义

1. 解析几何的基本思想及其意义

解析几何的产生，是数学由常量数学发展到变量数学的转折点，也是近代数学的开端．而解析几何的产生，通常以笛卡尔的"几何学"一文(发表于1637年的《方法论》一书的附录)为标志，和笛卡尔同时代的法国业余数学家费尔马也是解析几何的主要创始人之一．他在1929年写的《平面和立体的轨迹》一书就是用代数方法研究曲线的性质．因为费尔马这部著作是在他死后由其儿子整理出版的，所以人们往往忽略了他在创立解析几何上的贡献．

解析几何实现了代数与几何的有机结合．它的基本思想是：实数与直线上的点成一一对应，而实数对与平面上的点成一一对应；平面上的曲线可用含两个变量的代数方程来表示，而一个含两个变量的代数方程表示着平面上的曲线．实现这些基本思想的方法就是笛卡尔和费尔马提出的坐标方法．

解析几何的诞生对整个数学思想的发展产生了不可估量的影响，主要表现在以下几个方面：

(1) 变量思想开始进入数学，使数学思想方法发生了重大

的变革．

恩格斯曾经指出："数学中的转折点是笛卡尔的变数．有了变数，运动进入了数学．有了变数，辩证法进入了数学，有了变数，微分和积分就立刻成为必要的了"．可以说，17世纪以后的各种数学的发展都和变数联系在一起，变量思想进入数学使得数学能较好地解决工程技术及其他自然科学学科向数学提出的各种问题——主要是与运动变化有关的问题．变量思想已成为近代和现代数学中最重要、最基本的思想之一．

（2）提供了用代数计算解决几何问题的一种统一的思想方法．

古希腊时代的许多数学问题都是个别地解决的，而解析几何建立后就可以用代数方法作统一的处理．例如，解析几何产生后，平面几何中尺规作图的可能性才有统一的判定方法，从而使古希腊著名的三大尺规作图问题得以解决，圆锥曲线的性质研究才有可能统一于二次曲线的理论之中等等．

（3）促进了数学思想的新发展．

首先，许多数学概念得到进一步的抽象和概括．例如"曲线"概念，古希腊人只局限于能用简单的工具画出来的图形，而解析几何则把"曲线"理解为变量x、y的方程$F(x, y)=0$对应的几何图形，只要方程是由x、y的有限次代数运算构成即可，从而提出了研究"曲线"的新的思想方法——代数方法．更进一步，考察任意代数方程$F(x, y)=0$，研究它对应的曲线具有什么样的几何性质，由此发展出代数几何学的新思想．

其次，突破了几何直观的限制，开拓了发展数学的新思路，提出了新的数学思想方法．

例如，由实数与直线上的点一一对应及实数对与平面上的

点成一一对应，不难想到三元实数有序组与空间中的点成一一对应，再进一步自然要问四元，五元，…n 元实数有序组又对应于什么，由此导致线性代数(高维空间理论)的诞生．

(4) 将代数与几何结合起来，揭示了数学内在的统一性．

虽然欧几里得的《几何原本》中几何与算术(代数)是不加区分的，但他主要是用几何方法来处理各种数学问题．19世纪开始，代数学有了重大的发展，人们把代数与几何严格区分开来．这在当时对数学各科的深入发展是有利的，但数学毕竟是一个有内在联系的统一体，如果长期将数学中各分支割裂开来，不利于数学本身的发展．解析几何把几何与代数结合起来，用代数语言描述几何概念，用代数方法研究几何图形的性质以及给代数语言以几何解释，使代数概念变得直观易懂，易于理解，这种数形结合思想在近代和现代数学的发展中产生了深远的影响．

2．微积分的产生及其意义

变量数学发展的第二个重要阶段是微积分的创立．17世纪许多著名的数学家、天文学家和物理学家都参与了这方面的研究工作，其中贡献最大的要属牛顿和莱布尼兹．牛顿主要是从运动学来研究和建立微积分的，莱布尼兹则主要是从几何角度来研究和创立微积分的．

微积分思想的出现对数学思想方法的发展具有极其重要的作用，主要表现在以下几个方面：

(1) 微积分思想的出现，一方面，原有的常量数学的许多分支(如代数、几何、三角和数论)由于微积分思想的渗入，在内容上得到了极大的丰富，在思想方法上发生了深刻的变化．例如有些常量数学中的重要结论(如关于一元 n 次方程的根的代数基本定理)在借助于分析方法后才给出严格的证明．另一

方面，在微积分思想的影响下，新的数学分支如雨后春笋般地涌现出来．诸如常微分方程论、偏微分方程论、微分几何、复变函数论、解析数论等等．总之，微积分创立后，变量数学的思想方法很快就在整个数学的发展中占了主导地位，长期以来影响着数学发展的方向．

（2）微积分的产生，使自然科学描述现实物质世界的运动和变化过程的发展成为可能，扩大了数学思想方法的应用领域，从而促进了其他自然科学学科的进一步发展．

例如，18世纪的著名数学家欧拉不仅在数学领域(诸如微积分、解析几何、微分几何、微分方程、级数和变分法、数论等等)作出了卓越的贡献，而且他将微积分的思想方法应用到物理学的领域，创立了分析力学和刚体力学，并计算了行星轨道中的天体的摄动影响和阻尼介质中的弹道．他的潮汐理论、船舶航行和设计被用于指导航海，此外，他还用微积分的思想方法研究了梁的弯曲，计算了安全载荷等等．总之，由于自然科学的对象是运动变化的物质世界，而微积分是研究变量的数学，故微积分的思想方法必然成为研究物质世界的运动和变化规律的强有力的工具．

（3）微积分的思想方法为唯物辩证法的普遍规律在数学上提供了例证．

微积分中处处充满着矛盾：常量与变量、连续与间断、收敛与发散、微分与积分、有限与无限、近似与精确等等．微积分的思想方法充满了辩证思维．

二、随机思想的产生——概率论的创立及其意义

在人们的实践活动中常常遇到两类截然不同的现象：一类是必然现象，即这种现象在一定的条件下必然会发生或必然不

会发生．另一类现象是随机现象，即这种现象在一定条件下可能会发生某种结果，也可能不发生这种结果．

研究必然现象及其规律的数学称为必然数学，而研究随机现象及其规律的数学称为随机数学，从必然数学到随机数学是数学思想方法上又一次重大变革．

1．随机数学的产生

概率论是随机数学的基础理论，也是历史上最早出现的随机数学分支，它的产生同帕斯卡、费尔马和惠更斯的工作分不开．惠更斯专心研究了赌博中的计算问题，于1657年出版了《关于骰子游戏或赌博的计算》一书，成为了概率论诞生的标志之一．由于概率论很快在保险理论、人口统计、射击理论、年度预算、产品检验以及天文学和物理学中得到广泛的应用，因此，它从17世纪起一直成为许多数学家关注并作出许多成果的重要数学领域之一．

2．随机数学思想方法的意义

（1）随机数学思想方法的产生，使人们对现实社会中存在的数量关系的认识更深刻、更全面．

在现实世界中不仅存在大量的必然现象，而且存在大量的随机现象．虽然随机现象的发生不确定，但随机现象大量发生时，从总体上呈现出一定的规律．例如重复多次投掷一枚硬币，出现正面的次数约占总投掷次数的$\frac{1}{2}$．以概率论为基础理论的随机数学的诞生为现实世界中大量存在的这种统计规律的研究提供了一个强有力的工具．

（2）随机数学思想方法的产生拓广了数学在实践中的应用范围，进一步促进了数学思想方法的普及．

因为科学技术、国防、工农业生产中存在大量的随机现

象，所以以概率论为基础的随机数学诞生后，随着人们的深入研究，随机数学思想方法在各个领域内的应用越来越广泛，特别在电子技术、自动控制、气象预报、地震预报、洪水预报、地质勘探、企业管理、公共事业以及国防等部门的应用成果尤为丰富，随机数学的思想方法已深入到生产、生活和科学研究的各个领域．

（3）随机数学思想方法的产生促进了数学思想方法的新发展，产生了许多新的数学学科和分支，从而推动了整个数学的发展．

从18世纪末开始，约一个世纪的时间内，概率论中引入了母函数和特征函数的概念，并且引进了微积分等分析工具，使概率论进入了一个新的时期——分析概率论时期．到20世纪50年代，概率论的发展又进入了一个新的历史时期——现代概率时期，这个时期内概率论研究的重点是极限分布理论以及通过概率分布来研究随机过程，并且从50年代起，概率论形成了自己特有的思想方法——随机分析方法．在现代技术的刺激下，概率论的理论和应用都有很大的发展，分化为理论概率与应用概率两个大的方向．

现代概率论又形成了若干分支，如极限理论、平稳随机过程论、马尔科夫过程论、鞅和随机微分方程论、试验分析、多元分析等等．此外，随机数学的思想方法进入其他一些学科后，产生了不少边缘学科，如统计物理学、统计医学、统计生物学等等．随机数学的思想方法用于解决某类问题又产生了许多新的数学分支，如把它应用于公共服务事业而产生了“排队论”，应用于通讯技术而产生了“信息论”，应用于自动化生产而产生了“控制论”．

总之，随机数学思想方法的产生和发展标志着数学研究从

确定性事件领域深入到不确定性事件领域，不仅扩大了数学研究的范围，同时也是数学思想方法上的一次重大转折。

三、代数结构思想的确立——群论的诞生及其影响

1. 群论的诞生

群的概念是拉格朗日首先提出的。他在探讨代数方程的一般解法时提出了置换群的概念，形成了群论的初步思想。拉格朗日虽然给出了$n \leqslant 4$时一元n次方程有根式解，而$n \geqslant 5$时无根式解的基本道理，但并没有给出严格的证明。这一结论的严格证明是由阿贝尔给出的，阿贝尔的结论并不排除一些特殊的高次方程有根式解，但高次方程有无根式解的问题，是由伽罗华解决的。他引入了群论的思想，建立了现在称之为伽罗华可解性理论。在他的理论中引入了一系列新的数学概念：群、极大正规子群、可解群、伽罗华群等等。最早认识到群的价值并大力提倡的是英国的代数学家凯利和德国的戴德金。法国数学家约当于1870年出版的《置换和代数方程专论》成为群论的权威著作。在这部著作中，他系统地发展了有限群论，也用到现代群论的四条公理，建立了同构和同态等概念。约当的学生克莱因于1872年从变换群的观念出发，提出了著名的《爱尔兰根纲领》，将当时的各种几何学用到的变换群的思想进行了分类。约当的另一名学生于1874年引入了连续变换群(后来称之为局部李群)的概念，1883年他又引进了无限连续群的概念，主要用来阐明微分方程解的分类，他的理论思想深刻且内容广博。

2. 群论思想对数学发展的影响

(1) 群论的产生和发展使代数学的研究进入了一个新的时代，即“从局部性的研究转向系统结构的整体性分析研究的阶

段"(徐利治,《数学方法论选讲》)．群是最基本的代数结构，其他代数结构(例如环和域)可由群来确定．群论提供的结构分析思想是现代数学中最重要的思想之一．伽罗华把代数方程可解性问题化归为结构简单的问题的思想方法也成为现代数学的典型方法(徐利治,《数学方法论选讲》)．自从群论产生以后，研究各种结构的数学分支相继产生，如研究序结构的格论、研究拓扑结构的拓扑学、研究环和群的复合结构的模(model)论、研究同时具有几种结构的拓扑向量空间、微分流形、纤维丛等等．可以说结构思想是现代数学各分支中最基本、最重要的思想之一．

(2) 群论已成为解决数学难题的有力工具，并且群论促进了数学的进一步统一，例如代数方程的可解性问题、平面几何三大尺规作图不能问题等正是利用伽罗华理论解决的．也正是利用伽罗华理论，可以证明：如果 p 是形如 $2^{2^n}+1$ 的素数，则正 p 边形就可以用尺规作出．揭示了数论与几何之间的统一性．又如"同调群"既是拓扑学的重要概念，又是同调代数的基本概念之一．通过它将现代拓扑学与代数学统一起来了．由于群是各门数学分支中的一个基本概念，结构思想已渗透到数学各个分支，这就使人们更清晰地看到数学各分支间的内在联系．

(3) 群论思想扩大了数学思想方法的应用领域，促进了其他科学的发展．

四、模糊思想的产生——模糊数学的建立及其意义

1. 模糊数学的产生

20 世纪 60 年代产生了一门崭新的数学学科——模糊数学．它始于 1965 年美国自动控制论专家查德的开创性论文"模

糊集合”．从精确数学到模糊数学是数学思想方法的一个重大变革．

精确数学是建立在集合论的基础上．根据康托集合论的要求，一个对象对于一个集合而言，要么属于这一集合，要么不属于这一集合，二者必居其一且仅居其一，绝不允许模棱两可！但是就客观现实而言，大多数情形并不具有这种明晰性．如复杂系统的物理状态、生物学、医学、语言学以及其他一些自然科学和社会科学中出现的许多现象(高与矮、胖与瘦、美与丑、矿量的多与少、清洁与污染、长寿与短命等等)却无法用康托集合论来精确界定．这些模糊现象正是这些学科自身的特点所决定的．为了使这些学科的研究定量化和数学化，就必须引入新的数学思想方法，模糊数学正是在这样的背景下产生的．特别值得指出的是模糊数学的诞生和发展是与计算机的发展紧密相关的．因为计算机的一个发展方向是模拟人的大脑思维，而人的灵感思维和形象思维具有很大的模糊性，才使得人的大脑能灵活地进行思维．为了使计算机也能描述和处理事物的模糊性，从而完成更复杂的任务，就必须建立相应的能够描述和处理模糊量及其关系的数学思想方法．查德正是为了解决计算机发展中的这一矛盾而提出模糊集合这一概念的．自从查德成功地用数学方法描述模糊概念以来，模糊数学这门学科得到了迅速的发展并且广泛应用到各个领域．

2．模糊数学思想方法的意义

(1) 模糊数学思想的产生是数学思想方法的一次重大变革．

在经典的集合论中，一个命题或者为真，或者为假，二者必居其一．这种绝对的数学思想方法不能完全反映客观现实．特别在集合论中出现悖论后(即取真、取假都出现矛盾的命

题)，使这种绝对的数学思想方法产生了动摇．为了寻找解决的办法，导致了集合论公理化体系的出现，但这只是避开矛盾而没有解决矛盾．因而数学界对此仍有不同的看法，形成了不同的学派．查德并没有介入这些学派之间的争论，而是另辟蹊径，重新将模糊性和数学的精确性统一起来．查德并不是让数学放弃其严格性去迁就模糊性，而是让数学回过头来吸收人脑对于模糊现象识别和判断中的优点，引入隶属函数的概念，从而奠定了整个模糊数学的基础．由于模糊数学冲破了形而上学的束缚，既认识到事物的“非此即彼”的明晰性形态，又认识到事物的“亦此亦彼”的过渡性形态，因此，它较传统数学的应用领域更为广泛，从而具有强大的生命力．所以我们说从精确数学到模糊数学是数学思想方法上的一次重大变革．

（2）模糊数学思想的产生扩大了数学研究的领域，形成了新的数学研究方向．

模糊数学作为一门新兴的数学学科，虽然它的历史很短，但由于它是针对现代科学技术的迫切需要而产生的，因而对于它的基础理论的研究引起了广大数学工作者的注意和兴趣，它研究的课题越来越广泛．人们试图把经典数学的概念和结论都推广到模糊数学中去，经过众多数学工作者的努力，目前已取得丰硕成果．它们涉及的范围有模糊数、模糊关系、模糊矩阵、模糊图、模糊映射和变换、模糊概率、模糊规划、模糊逻辑、模糊识别和模糊控制等等，形成了模糊分析学、模糊拓扑学、模糊概率论、模糊语言学等新的数学研究方向．

（3）模糊数学思想的产生扩大了数学思想方法的应用范围，促进了其他学科的进一步发展．

近 30 多年来，模糊数学的思想方法已广泛渗透到科学和技术的各个领域．例如在医疗诊断中，一些症状如“食欲不

振"、"精神不好"、"头痛"等都是模糊概念，而某种病症实际上是由一些症状所组成的模糊集合，因此，编制计算机诊断程序时，采用模糊数学模型更符合客观实际，效果会更好些，除了医疗诊断外，模糊数学在天气预报、林业、农业、建筑、采矿、心理、生物、冶金、地质、人工智能等领域内找到了广泛的应用，为解决科学技术各部门的一些难题，提供了一种新的思想方法．

五、非标准思想的创立——非标准分析的建立及其意义

1．分析的严格化

前面我们已经讲到，微积分创立初期带有许多神秘的色彩，到了19世纪初，人们针对微积分的基础理论提出了一系列需要解决的问题，其中最重要的是：无穷小量是什么？怎样建立实数和实数系统？函数连续的严格定义和无穷级数的收敛定义是什么？等等．为了解决上述问题，法国数学家柯西首先给出变量极限的定义，同时用极限给出无穷小的定义，对于无穷大，柯西将其看成无限地变大，其值可以超过任何固定常量的变量，再利用极限给出导数和积分的定义，进而又通过导数定义微分和高阶微分，从而给出微积分基本定理的精确叙述和严格证明．总之，柯西用极限理论对牛顿和莱布尼兹的微积分进行了彻底地改造，这是一个重大的进步．但柯西的体系仍值得进一步改进．例如，柯西没有一致收敛的概念，因而有的定理无法证明，有的定理的证明不够严格，甚至柯西有些判断后来证明是错误的．更主要的是柯西的极限概念中含有一些不确切的语句，如"越来越近"、"无限地减小"、"无限地趋近"等，都不是数学语言．稍后德国数学家魏尔斯特拉斯对此作了重大

的改进．他提出了同已有的变量和极限的动态观点完全不同的静态观点．也就是把柯西关于极限的定性描述改造成定量的叙述，即现代数学分析中所采用的“$\varepsilon-\delta$”定义．最后，由戴德金、康托等数学家完成了实数系和实数理论的建立工作，从而使数学分析奠定在严格的理论基础之上．

2．非标准分析的创立

19 世纪末，在极限和实数理论基础上建立的数学分析，现在称之为标准分析．在标准分析中，函数 $f(x)$ 在一点的极限和导数必须由 $f(x)$ 在此点邻近的值的变化状态来确定．也就是只承认点与点之间的函数值的联系（外部联系）而不涉及该点本身的内部矛盾．所以标准分析中的极限思想虽然消除了微积分的神秘性，但仍然没有摆脱形而上学的束缚．在标准分析的基本概念中蕴含了形而上学观点的原因是它的论域中排除了无穷小量，因而标准分析中的点只能是既无内部结构、又无层次的抽象点．

本世纪以来，一些数理逻辑学家为了摆脱形而上学对数学分析的束缚，不用极限方法，而用数理逻辑方法去建立数学分析的理论体系．1960 年，美国数理逻辑学家鲁宾逊用数理逻辑的模型论的思想证明了无穷小量的存在，把无穷小和无穷大纳入了实数域而形成了超实数域，并在此基础上，建立起全部数学分析理论，人们称之为非标准分析．

3．非标准分析的意义

（1）非标准分析的产生揭示了数学内在的统一性．人们素来认为像数理逻辑这样非计量科学与微积分之间的联系不大，但非标准分析揭示了它们之间的内在联系．这说明数学各分支间联系日趋加强，这也是现代数学发展的必然趋势．因此，我们必须十分重视一个数学分支的成果和方法在其他分支中的应

用．很多数学难题长期不能解决是因为缺乏有力的工具和有效的方法，而这些工具和方法常常不是在传统的探索范围内产生，而是来自其他领域．例如数论中的著名难题——费尔马大定理，最终被证明恰好是借助于代数几何中的结论和方法．

（2）非标准分析的产生证明了数学发展的相对独立性，同时也说明只有敢于冲破理论禁区，才能促进数学向前发展．

数学理论的发展虽然依赖于社会实践，但又有其相对独立性．事实上，当数学发展到一定阶段后，往往会出现一些矛盾，这时需要从理论上加以提炼和概括，而解决这些矛盾的过程往往导致数学理论自身相对独立的发展．这种相对独立的发展是依赖于正确的前提条件和正确的逻辑方法的．

在标准分析中点是不可分的，标准实数轴上的点是一个"只有位置、没有大小、没有结构"的抽象的点．这样"点的结构"就成了一个理论禁区．于是在标准分析中，不同阶的微分所表现出不同阶的无穷小量在实数轴上无法直观地表现出来，从而使它很抽象，难于理解．而非标准分析打破了这一传统观念，引入了"非标准实数"的概念，认为"点有结构"，建立了单子结构模型，把实数域扩充到超实数域，这样便能把不同阶的微分，清楚地、直观地展示在非标准实数轴的不同层次的单子里面(每一个单子里只含一个标准实数，其余都是无穷小量，并且称这些无穷小量为非标准实数)，从而使标准分析产生了新的飞跃．于是，牛顿—莱布尼兹的初级微积分直观概念被柯西—魏尔斯特拉斯等人的严密、抽象的极限理论所扬弃，而柯西—魏尔斯特拉斯等人的极限理论又被鲁宾逊的非标准分析所扬弃，这种"直观—抽象—直观"的历史演变过程，恰好符合辩证法的规律，从而也说明非标准分析是微积分理论自身发展的一个必然产物．

(3) 非标准化思想是一种重要的数学思想.

从数学发展历史的各个阶段来看，非标准化思想对数学的发展起着重大的作用．例如，变量对常量而言是非标准的；非欧几何对欧氏几何而言是非标准的；模糊集合论对康托集合论而言也是非标准的．可见，数学理论由“标准”向“非标准”转化，是解决数学理论体系内部矛盾，推动数学向前发展的一个重要途径．在数学中自觉实现“非标准化”是数学思想方法上的一个重大变革.

总之，自觉实现由“标准”向“非标准”的转化，必将促进数学理论相对独立的发展.

§2.5 电子计算机对数学思想发展的影响

从50年代开始，电子计算机得到了迅速发展，经历了电子管计算机、晶体管计算机、集成电路计算机、大规模集成电路计算机和并行计算机五个时代，它的应用已经遍及科学技术的每一个领域．电子计算机的产生，不仅是本世纪最伟大的发明之一，而且也是迄今为止人类历史上最广泛、最深刻的技术革命.

电子计算机的产生对数学思想方法的发展带来深刻的影响，主要表现在下列几个方面：

一、促进了数学思想方法的普遍化

由于电子计算机能代替人的一部分脑力劳动，使科学研究(包括数学研究)的方式产生了深刻的变化．科学研究方式的变化反过来又影响了科学的发展．各种科学研究都在利用电子计算机这一先进工具，从而使各种科学研究的发展趋势更加数学

化．一方面，人们在用电子计算机解决某个领域内的问题时，首先应将这个问题形式化，即建立这个问题的一个数学模型；其次要设计该模型问题的一个算法；此外，还要讨论计算的复杂性及计算误差，即计算机能否在可接受的时间内完成设计的计算以及计算产生的误差是否在允许的范围内．要解决这些问题，就必须采用数学思想方法(如数学模型法、公理法等等)，从而促进数学思想方法应用的普遍化．另一方面，由于新一代计算机的研制涉及到计算、证明等数学思维的本质问题，对这些问题的研究与探讨，促使数学向更为抽象的方向发展，从而使数学思想方法具有更大的普遍适用性．

二、形成了新的数学学科和新的数学理论

首先，由于电子计算机的研制和应用，形成并丰富了与之相关的一系列数学理论，如计算方法、计算机数学、计算可行性和复杂性理论以及并行算法理论等等．由于电子计算机的发明和发展，计算方法已成为与理论方法和实验方法并列的重要科学研究方法．运用算法去解决问题离不开数学模型法和数学推理，因此，与计算机相关的一系列数学理论的形成与发展是各种数学思想方法综合运用的结果．

其次，计算机的发明和发展，对科学的发展和科学结构的形成产生了很大的影响．计算方法与其他数学分支相结合产生了一系列新的数学学科，如计算几何、计算概率、计算统计、计算数论、数值代数、微分方程的数值解、计算机绘图等等．此外，计算方法与其他科学相结合形成了一系列边缘科学和交叉学科，如计算力学、计算物理、计算化学、计算生物学、计算地震学、计算地质学、数值气象学等等．由于这些新的学科的建立，形成了一系列新的理论和新方法．例如，计算几何

是由函数逼近论、微分几何、代数几何、计算数学等结合起来形成的新学科，它研究几何外形信息的计算机表示、分析和综合，它是计算机辅助几何设计（CAGD）这门技术的数学理论基础．从20世纪60年代起，在造船、飞机设计、汽车工业等几何外形设计和制造中有广泛的应用．这些实际应用反过来又提出了一系列新的数学问题，这些问题的解决又形成了新的理论．例如为了保证曲线和曲面的光滑、凸凹，形成并发展了新的插值理论——样条插值理论．

三、促进了数学思想方法的新发展

首先，电子计算机的诞生使数学中采用的两种基本方法——证明和计算趋于统一．方法上的统一性揭示了数学本身的统一性．一方面利用电子计算机可以通过计算来证明数学中许多不平凡的定理．例如20世纪50年代末，美籍华人数学家王浩发明了王浩算法，把机器过程规则化．1958年他使用电子计算机只用了3分钟的时间证明了《数学原理》中约200个命题，次年他又用不到4分钟的时间证明了另外133个一阶逻辑运算定理．1956年美国数学家鲁宾逊提出了归结原理，把机器证明向前推进了一大步．接着非归结原理的机器证明方法也有很大的进展．我国著名数学家吴文俊从1977年起陆续提出了初等几何和初等微分几何的判定原理，不仅用电子计算机证明了人们难以证明的问题，而且还发现了一批新的平面几何定理．我国数学家还在吴氏机械化数学理论的基础上提出了以检验数值实例为基本手段的机器证明方法——洪加威1986年提出单点例证法，张景中和杨路于1988年前后提出的数值平行法，也就是说只要用电子计算机对一个或几个满足一定条件的特殊例子加以计算验证，就可作出一般性的结论．这两种新的

机器证明方法所揭示的数学思想是深刻的．单点例证法显示了从特殊事物到一般事物的联系，而数值并行法则揭示了大量经验中必蕴含着客观规律的奥秘．数值并行法说明了在一定条件和范围内，经验归纳法(即不完全归纳法)得出的结论同样是正确的．这对传统的数学思想——经验归纳法得出的结论不一定正确——不能不说是一个很大的挑战．机器证明另一个成功的例子是1976年美国的阿佩尔和哈肯等人在电子计算机上解决了100多年悬而未决的数学难题——四色猜想，肯定地证明了它．又如，美国著名的数值分析专家瓦格于1990年出版了《数学猜想和科学计算》一书，作者利用高精度的计算工具及符号运算工具去解决若干纯数学问题，如多项式逼近论中的伯恩斯坦猜想、解析数论中黎曼猜想及相关的一些问题等等，均取得了惊人的进展．这些事实雄辩地证明了电子计算机不仅可以成为工程师和应用科学家的得力助手，还完全可以成为纯数学家的强有力的工具．另一方面，由定理的证明也可以导出算法．例如1974年美国的几位计算数学专家提出了求映射不动点的著名的连续同伦算法，这一算法就是在赫尔给出的著名的不动点定理的一个证明方法的基础上提出的．可见人们不仅可用计算来进行证明，反过来也可通过证明来研究和导出算法．人们通常总认为计算是比较单调、复杂而工作量大的工作，而证明是比较困难、技巧性强的高质量的智力劳动．机器证明则使两者结合起来，用大量的、复杂的单调工作取代高质量的、技巧性强的困难工作，充分反映了数学中计算和证明这两种思想方法在本质上的一致性．

其次，电子计算机的发展使得离散化思想方法得到进一步发展，从而促进了处理离散对象的数学的发展．因为数字电子计算机本质上是离散的，它的特点是根据有限多个有限位数据

进行有限次运算，算出有限个有限位的解答来．而生产实践、科学研究及日常生活中遇到的许多问题却是连续性的(如求非线性方程(组)的解以及求微分方程的解等等)，为了用计算机解连续性问题，首先应将这些问题离散化(例如将微分方程差分化变为代数方程)，然后用电子计算机得出离散解答来．因此，计算机的出现导致了离散化数学思想的广泛应用，具体说来，也就是导致了研究离散结构的各数学分支(如数理逻辑、图论、组合数学、代数系统等)的综合应用．并且，新型计算机的设计、制造和使用又向数学本身提出许多新课题，为了研究解决这些新问题，导致了对这些数学分支的深入研究，从而促进了这些数学分支的进一步发展．

最后，电子计算机的出现使数学模拟思想方法的应用领域更为广阔，成为了人们解决问题的一种重要方法．模拟方法是一种根据原型(研究对象)的某些特性、关系和特征，人为地建立一个与之相似的模型，并通过模型来研究原型，它实质上是一种类比方法．数学模拟方法主要有二种：一种是利用某些物理量之间的关系与某种数学表达式相似从而用这种物理量的变化来模拟运算．这种模拟过程通常在模拟电子计算机上完成，模拟电子计算机的优点是计算速度快且计算速度可按人的意图加以改变．此外，模拟电子计算机可以与实物连接进行局部模拟，便于对实际过程进行实时控制，再加上模拟电子计算机结构较简单、成本较低，从而使模拟电子计算机的应用十分广泛，它主要用于科学计算(如解各种微分方程)、对复杂自然现象或工程目标进行模拟(如模拟破坏性实验，得出设计强度的要求，又避免了价格昂贵的破坏性实验)、构成控制系统的动态模拟，以及作为模拟训练器，用来训练飞行员、宇航员、远洋轮船驾驶员、坦克手等等．另一种模拟方法是数学模拟方

法，即把数学模型数字化，用数字来模拟被模拟对象，用数字模拟法来研究对象的性质和特征，数字模拟过程通常在数字电子计算机上完成，它被广泛地应用到各种领域．例如，进行空气动力学的模拟计算以代替风洞试验；模拟各种粒子反应以代替费用高昂的高能物理实验．模拟化学反应，不进行化学实验而得出其结果；模拟爆炸过程，而不必进行破坏性及危险性极大的实际爆炸等等．

总之，由于模拟电子计算机及数字电子计算机的产生，使数学模拟思想方法成为了人们解决实际问题的一种重要的思想方法．

第三章　数学中的两大基本思想

§3.1　两种基本数学思想

在前面两章中，我们对数学思想方法的本质属性(内涵)及其发展历史作了简要的分析，这样就为我们论述具体的数学思想(外延)作了充分的准备．

构造思想和转化思想是数学中的两大基本思想，这是由数学和数学方法的本质所决定的．

恩格斯曾经指出："数学的对象是现实世界的空间形式和数量关系，这是非常现实的材料．"存在于现实世界的这种"形式"、"关系"具有结构的性质，它们是数学研究的对象．恩格斯又说："这些材料以极度抽象的形式出现．""为了能够从纯粹的状态中研究这些形式的关系，必须使它们完全脱离自己的内容，把内容作为无关重要的东西放在一边．"因此，数学研究的是纯粹的、极度抽象的、完全撇开具体内容的形式和关系．也就是说，数学的对象具有明显的纯粹结构性质的特点，甚至有一种观点认为，数学的对象是结构．

以布尔巴基为代表的结构主义学派这样地描述数学："数学好比一座大城市，城市中心有些巨大的建筑物，就好比是一个个已经建成的数学理论体系．城市的郊区正在不断地并且多少有点杂乱无章地向外伸展，它们就好像是一些尚未发育成型的正在成长着的数学新分支．与此同时，市中心又在时时重建，

每次都是根据构思更加清晰地计划和更加合理地布局，在拆毁掉旧的迷宫似的小街小巷的同时，将修筑起新的更直、更宽、更方便的林阴大道通向四方…”，这就是说，数学亦如城市，是一个逐步地建造起来的结构．

数学的这种结构性的特点决定了构造是数学中的一种基本思想．我们常说的“列方程”、“作图”、“建立坐标系”、“建立模型”、“构造算法”、“建立公理体系”等等，都具有明显的构造性的色彩．许多数学问题的求解，当我们把具体的对象构造出来以后，问题也就完全解决了．这比千言万语的证明更为有效．

勾股定理的证明据说已有数百种，印度数学家婆什迦罗的一个证明是画了一个如图 3－1 的图形，他只写了一个字：“瞧！”

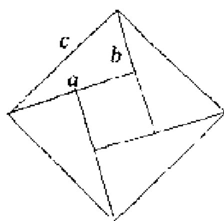


图 3－1

这个图显然来自中国的“弦图”．它把要证明的东西都构造出来了：

$$c^2=4\left(\frac{ab}{2}\right)+(b-a)^2=a^2+b^2 .$$

要证明一个大数是合数还是质数，有时是很困难的．数学家们在很长的一段时期内曾希望知道 $2^{67}-1$ 是否为质数．但直到 1903 年，在美国纽约举行的一个学术报告会上，数学家科尔走上讲台，一言不发，只在黑板上写了一行字：

$$\begin{aligned}2^{67}-1 &= 147\ 573\ 952\ 589\ 676\ 412\ 927\\ &= 193\ 707\ 721\times 761\ 838\ 257\ 287 .\end{aligned}$$

听众席上马上爆发出热烈的掌声，祝贺他的成功．因为他把这个合数的结构构造出来了．

如果说构造方法是数学中的直接方法，那么转化就是数学

中的间接方法．

转化也是数学中特有的思想．结构 A 和结构 B 尽管具体内容不同，但从抽象的关系来看其本质可以相同或部分相同．这样，在结构 A 中陈述的问题可以在结构 B 中找到相应的陈述，反之亦然．如果在结构 A 中问题所涉及的关系比较隐晦而难于处理，而在结构 B 中相应的问题反倒比较明显而易于处理，则我们将 A 中的问题转化为 B 中的问题显然有助于问题的解决．由此可见，正是数学的高度抽象的结构性的特点为转化思想提供了前提和基础．

匈牙利数学家罗莎曾经对转化作过十分形象的描述：假设在你面前有煤气灶、水笼头、水壶和火柴，现在的任务是要烧一壶水，你应该怎样做？这个问题很简单，谁都知道做法：先拿水壶到水笼头下装满水，再划火柴点燃煤气，然后把水壶放到煤气灶上．接着罗莎再提出一个问题：假设所有的条件仍和原来一样，只是在水壶中已装满了水，这时你又应该怎样去做？对于第二个问题人们往往会这样回答：那就只要划火柴点燃煤气，再把水壶放到煤气灶上．但是罗莎指出：对数学家来说，这不是最好的回答，因为只有物理学家才会这样做，而数学家则会倒去壶中的水，并且声称已经把第二个问题转化成第一个问题，而第一个问题是已经解决了的．

罗莎的比喻当然带有十分夸张的成分，作为生活中的事件，这个例子你也许觉得荒诞，但在数学中我们却经常不断地这样做：我们总是不断地将新问题化归为已解决的问题，并以此宣布新问题的解决．甚至有许许多多等价的问题，只要解决其中的一个，则所有这些问题全部解决．化归方法体现了数学研究的继承性，正是前辈数学家积累了许许多多已解决的问题并且建立了严谨的理论，化归才成为可能．每个数学家都站在

他的前辈的肩膀上推进数学，数学的发展才有今天的局面．

构造和转化分别作为数学方法中的直接方法和间接方法，两者互相补充．从结构论的观点来看，构造是转化的基础．大量的例子表明，合适的构造往往又是促成转化的手段．正是因为构造了直角坐标系，几何问题才得以转化为代数问题，从而可以用代数的方法来解决．不同结构之间的转化又体现了各种结构之间的相互联系和内在关系，转化使其相互沟通．所以构造和转化之间有着十分密切的关系．

现在，我们来看两个转化与构造互相配合的例子．

例 1 已知$\triangle ABC$的外接圆与内切圆半径分别为R和r．求证：$\frac{r}{R}\leqslant\frac{1}{2}$．

我们熟知两相似三角形的内切圆或外接圆的半径之比就等于相似比．如果我们能把r与R都变成两个相似比小于$\frac{1}{2}$的相似三角形的内切圆（或外接圆）的半径，则命题得证．为了完成这一转化，我们必须构造出两个相似三角形$A_1B_1C_1$和$A_2B_2C_2$，使得$\triangle A_2B_2C_2$以$\triangle ABC$的外接圆为内切圆，即使其内切圆半径为R，而使$\triangle A_1B_1C_1$的内切圆半径为$2r$，且使$\frac{A_1B_1}{A_2B_2}\leqslant 1$．一个具体的构造办法见图3－2．

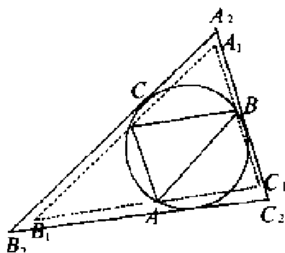


图3－2

例 2 试求定义在自然数集上的函数$A(n)$，使满足条件：

$$\begin{cases}A(1)=1,\\ A(m+n)=A(m)+A(n)+mn.\end{cases} \quad ①$$

这是一个典型的需要构造的问题，即构造出一个合乎定义的函数$A(n)$. 但由于①中第二个等式中有两个变量，不好控制. 我们不妨令 $m=1$，先把它转化为

$$\begin{cases} A(1)=1, \\ A(1+n)=A(n)+n+1. \end{cases} \quad ②$$

②显然好构造得多，顺次令 $n=1, 2, \cdots, n$，即得：

$$A(1+1)=A(2)=A(1)+2,$$
$$A(1+2)=A(3)=A(2)+3,$$
$$\cdots$$
$$A(1+n)=A(1+n)=A(n)+(n+1).$$

将 n 个等式相加，即得

$$A(1+n)=A(1)+2+3+\cdots+(n+1)$$
$$=\frac{1}{2}(n+1)(n+2). \quad ③$$

于是 $A(n)=A[1+(n-1)]=\frac{1}{2}[(n-1)+1][(n-1)+2]$

$$=\frac{1}{2}n(n+1).$$

这样就得出所要构造的函数了.

§3.2 转化思想概论

一、转化的作用

转化是把待解决的问题从一种形式转化为另一种形式，使之较易于解决. 转化的作用十分广泛，并不局限于个别具体问题的解决.

1. 优化解题方法

追求解题方法的简捷、深刻、优美，是数学思想的最大特色．有些数学问题通过转化，不只是获得了解决，更重要是获得了解法的优化．

著名数学家华罗庚教授曾介绍过一个有趣的猜帽问题：

有一位老师，想辨别一下他的三个得意门生中哪一位更聪明一些，他采用了以下的方法：事先准备5顶帽子，其中3顶是白的，2顶是黑的．试验时，他先把这些帽子让学生看了看，然后要他们闭上眼睛，替每个学生戴上一顶白色的帽子，并且把2顶黑帽子藏了起来，最后让他们张开眼睛，说出自己头上的帽子，究竟是哪一种颜色．

三个学生相互看了一看，踌躇了一会儿，然后他们异口同声地说自己头上戴的是白色的帽子．

他们是怎样推算出来的呢？下面给出一种解答：先不考虑三个人，而转化为仅仅考虑二个人1顶黑帽子的问题．这个问题很容易解决，因为黑帽子只有1顶，如果我戴了，那么另一个人会很快说出自己戴的是白帽子．但是，他为什么踌躇呢？可见我戴的不是黑帽子而是白帽子．

现在来看三个人的情况，“2顶黑帽子，不管多少顶(当然不小于3)白帽子”的问题就很容易了．为什么呢？如果我头上戴的是黑帽子，那么对另外二人来说就变为“二个人，1顶黑帽子”问题，他们就会立刻回答“自己头上戴的是白帽子”而不必踌躇．现在他们踌躇，正说明了我头上戴的是白帽子而不是黑帽子．

这个问题经过这样一转化，不仅可以推广到多于3个学生的情况，而且把问题的本质揭露出来了．

2．揭露问题本质

历史上有不少数学问题，在原来提出这一问题的领域内很

难解决，甚至无法解决，就像人不能自举其身一样。"不识庐山真面目，只缘身在此山中"，如果把问题转化到另一领域中，就可以迎刃而解了。例如，著名的古希腊几何作图三大难题，在欧氏几何中长期未能解决，直到上世纪，把它转化为代数问题后才彻底解决。

古希腊人留下的几何作图三大难题是：

1.（三分角问题）将一个任意角三等分；

2.（化圆为方问题）求作一个正方形使其面积等于已知圆的面积；

3.（倍立方问题）求作一立方体，使其体积等于已知立方体体积的 2 倍。

在这里作图的工具只限于圆规和无刻度的直尺。

这三个问题早在公元前 4 世纪就已流传，这三个问题貌似平凡简单，实则隐含玄机，历史上曾有不少学者为之耗费了大量精力和时间，但在欧几里得几何中却始终得不到解决，直到 19 世纪，人们利用解析几何与近世代数的知识，才找出这三大难题的答案。

因为上述三个问题的作图，归根结底是要作出一线段，使其长度等于已知数，要作一线段等于已知的线段，当线段的一个端点确定之后，问题就转化为确定另一个端点的位置，解析几何与代数方法得出的结论是：

由已知的长度使用圆规与直尺可作出的新长度都可以由已知的长度用有限多次的加减乘除及开平方得到；反之，如果一个长度不能由已知的长度经过有限次加减乘除及开平方得到，则此长度不能用尺规作出。

现在可以说明古希腊几何作图三大难题为什么不可能了。例如三分角问题，根据三角恒等式：

$$\cos 3x = 4\cos^3 x - 3\cos x .$$

令 $x=20°$，并记 $\cos 20° = y$，则得三次方程

$$4y^3 - 3y - \frac{1}{2} = 0 .$$

根据三次方程的求根公式知此方程的惟一正实根必须通过方程系数的立方根表示出来．所以 $y=\cos 20°$，不可能用尺规作图．这就意味着 60°的角不能用尺规三等分．特殊的角 60°尚且不能三等分，对于任意角就更不必说了．

3．开辟研究领域

在区组设计等问题中，要用到所谓"正交拉丁方"．正交拉丁方这一重要的数学分支，来源于一个有趣的传说．

18 世纪时，普鲁士的腓特烈大帝要举行一次盛大的阅兵式，他心血来潮突然想到要组成一个由 36 名军官组成的方阵．普鲁士有 6 种部队，每一种部队有 6 种不同的官阶．腓特烈大帝要求每种部队选派 6 名不同级别的军官，共 36 人排成一个 6×6 的方阵，要求每一行每一列都有各部队、各官阶的军官．腓特烈大帝一声令下，可忙坏了阅兵司令，他按要求选来了 36 名军官，但左排右排，怎么也排不出腓特烈大帝所要求的方阵．腓特烈大帝亲临指挥也无济于事．事后，他们只好去请教当时的大数学家欧拉，这个问题就被称为"36 军官问题"．

欧拉按照通常的数学思想，从特殊到一般，以便从中发现规律．他用大写拉丁字母 A，B，C…表示军种，小写拉丁字母 a，b，c…表示官阶，他很快排出了 3×3 (3 个军种，3 种官阶)、4×4 (4 个军种，4 种官阶) 和 5×5 (5 个军种，5 种官阶)的方阵．例如，下面是一个 5×5 的方阵：

Aa	*Bb*	*Cc*	*Dd*	*Ee*
Ed	*Ae*	*Ba*	*Cb*	*Dc*
Db	*Ec*	*Ad*	*Be*	*Ca*
Ce	*Da*	*Eb*	*Ac*	*Bd*
Bc	*Cd*	*De*	*Ea*	*Ab*

当 $n=2$ 时，2×2 的方阵显然排不出来，对于 $n=6$，连欧拉这样的大数学家也没有排出来，于是他提出一个猜想：若

$$n\equiv2(\mathrm{mod}4),$$

则 $n\times n$ 方阵排不出来．这时欧拉年事已高，来不及证明这个结论便于第二年去世了．

这个看似简单的问题由于大数学家欧拉的关注而引起数学家的兴趣．数学家们认识到，必须找出一种研究方法来处理这类问题，而不能靠枚举的方法来解决．于是人们把腓特烈大帝所要求的那种方阵称为欧拉方阵．一个欧拉方阵可分解为两个方阵，一个由大写字母组成，一个由小写字母组成．为了方便，人们把小写字母的方阵也改写为大写字母的方阵．这样由字母组成的方阵称为拉丁方．若两个拉丁方能够重合成一个欧拉方阵，则称它们是正交的．

于是，欧拉猜想就转化成：

当 $n\equiv2(\mathrm{mod}4)$时，不存在 n 阶正交拉丁方．

遗憾的是，“智者千虑，必有一失”，欧拉的这一猜想，竟然只有 $n=2$，$n=6$ 时是正确的，实在出人意料．但有幸的是：当人们把腓特烈大帝所要求的方阵映射为拉丁方以后，便产生了区组设计这一新的数学分支．

二、转化的方向

转化的基本目的是把待解决的问题转化为易于解决的问

题，但怎样才能达到这一目的呢？那却只能具体问题具体分析，而不可能有万灵的药方，下面介绍在初等数学中常见的几种转化方向．

1．转化为特殊情况

有的数学问题所要求的结论，在一般情况下不容易推出，但在特殊情况下非常易于处理，并且在很多时候特例对一般情况的解决有奠基或桥梁的作用，因此把一般问题转化为其特例，常有助于问题的解决．

例 求证：

面积等于 1 的三角形不能被面积小于 2 的平行四边形所覆盖．

首先我们把问题转化为与它等价的逆否命题：

1）若$\triangle PQR$的顶点在$\square ABCD$的内部(包括在边界上)，则$S_{\triangle PQR}\leqslant\frac{1}{2}S_{\square ABCD}$．

把问题转化为1)的形式之后，结论更显得直观．在图 3－3 中，对于图形(a)那样的特殊情况，即$\triangle PQR$有两个顶点在平行四边形的同一边上时，显然有

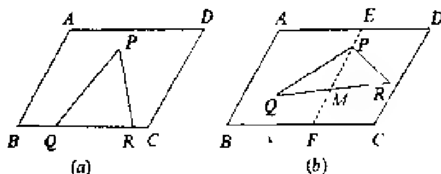


图 3－3

$$S_{\triangle PQR}\leqslant\frac{1}{2}S_{\square ABCD}.\qquad ①$$

对于图(b)那样的一般情况，只要通过$\triangle PQR$的任一顶点例如P，引平行于AB的直线，分别交AD、QR、BC于

E，M，F．则转化为两个图(a)那样的特例了．这时有

$$S_{\triangle PQR}=S_{\triangle PQM}+S_{\triangle PMR}$$
$$\leqslant\frac{1}{2}S_{\square ABFE}+\frac{1}{2}S_{\square EFCD}$$
$$=\frac{1}{2}S_{\square ABCD}.$$

2．转化为奠基条件

这一类问题的特征是，问题甲在问题乙的基础上求解，问题乙又在问题丙的基础上求解．这样逐步转化下去，一直追溯到最原始的基础，然后逆其次序即可得到问题的解．

例 求自然数的方幂和．设$f_k(n)$表示前n个自然数的k次幂之和，即

$$f_k(n)=1^k+2^k+\cdots+(n-1)^k+n^k.$$

这是一个很令人感兴趣的问题，已有多种解法，但绝大多数解法都建立在共同的思想上，即都依赖递推公式．欲求$f_k(n)$可转化为求$f_{k-1}(n)$，$f_{k-2}(n)$，…，$f_2(n)$，$f_1(n)$，再反过来利用$f_1(n)$，$f_2(n)$，…，$f_{k-1}(n)$的结果表示出$f_k(n)$．

试以求$f_3(n)$的一种方法为例．

$$f_3(n)=1^3+2^3+\cdots+n^3,$$

利用恒等式

$$(n+1)^4=n^4+4n^3+6n^2+4n+1.$$

顺次令$n=0$，1，2，…，n，则得

$$1^4=1,$$
$$2^4=1+4\cdot1+6\cdot1^2+4\cdot1^3+1^4,$$
$$3^4=1+4\cdot2+6\cdot2^2+4\cdot2^3+2^4,$$
$$\cdots$$
$$(n-1)^4=1+4(n-2)+6(n-2)^2+4(n-2)^3$$

$$+(n-2)^4,$$

$$n^4 = 1+4(n-1)+6(n-1)^2+4(n-1)^3+(n-1)^4,$$

$$(n+1)^4 = 1+4n+6n^2+4n^3+n^4.$$

相加即得：

$$(n+1)^4 = 4f_3(n)+6f_2(n)+4f_1(n)+(n+1).$$

利用这一恒等式就把求 $f_3(n)$ 的问题转化为求 $f_2(n)$ 和 $f_1(n)$ 的问题，下一步又可以把 $f_2(n)$ 转化为求 $f_1(n)$ 的问题，如果 $f_1(n)$ 是已知的，则 $f_3(n)$ 也就解决了。

数学中常用的递推、迭代、数学归纳法、分析法(倒推法)等等都是运用这一类转化思想。

3. 转化为典型状态

转化为典型状态也是数学中最常见的转化方向，所谓典型问题是指那些具有标准形式和固定解法程序的问题。例如一元二次方程有固定的解法，我们就常把一些无理方程、分式方程等转化为标准的二次方程求解。等比级数有固定的求和方法，我们在求级数的和时，就看它能否化为等比级数以求和。其他诸如数学中的换元、变换等，大都也是把求解的问题转化为有固定解法的标准状态。

例 有5只猴子分一堆桃子，它们想平均分配，但无论如何也分不下去。于是大家先约定去睡觉，明天再分。晚上第一个猴子趁大家熟睡，走到桃子边，先吃掉一个后，剩下的恰好可以平均分成5份。这个猴子把其中一份藏了起来，然后重新去睡觉。过了一会第二个猴子起来，在剩下的桃子中拿一个吃了，其余的又恰好能平均分成5份。第二只猴子也藏了其中一份后去睡觉。接着第三、第四个猴子都照此办理。最后第五个猴子起来，把剩下的桃子吃掉一个以后，余下的桃子也恰能平均分成5份。问最少有多少只桃子。

这是一个有名的趣题，曾引起许多数学家的注意．它有多种解法．开始，人们大多用不定方程求解，比较麻烦．特别是如果把5只猴子推广为 n 只猴子，计算量就更大．利用转化思想，把它化为一个典型的等比数列问题，就可以得到一个较简捷的解法：

设开始有 x 个桃子，n 只猴子．第 k 只猴子藏去的一份桃子数为 x_k，则得到一个正整数列：

$$x_1,\ x_2,\ \cdots,\ x_n$$

它的项满足递推关系：

$$(n-1)x_{k-1}-1=nx_k,$$

两边加上 n，则得

$$(n-1)(x_{k-1}+1)=n(x_k+1).$$

令 $y_k=x_k+1$，则 $y_1,\ y_2,\ \cdots,\ y_n$ 是一个公比为$\frac{n-1}{n}$的等比级数：

$$y_k=\frac{n-1}{n}y_{k-1},$$

或
$$y_n=\left(\frac{n-1}{n}\right)^{n-1}y_1.$$

由于 y_n 是正整数，y_1 最小应为 n^{n-1}，从而 x_1 最小为 $n^{n-1}-1$．再注意到

$$x-1=nx_1,$$

即得 x 的最小值为 $nx_1+1=n(n^{n-1}-1)+1=n^n-(n-1)$．

在本题中取 $n=5$，即得 $x=5^5-4=3121$．

4．转化为渐近过程

在数学研究中，许多问题常常不可能一开始就直接获得解决，往往退而求其次，采取迂回包抄、逐步逼近的办法求解．因此，常把问题转化为渐近状态．

著名的世界级难题"哥德巴赫猜想"的研究，就采取了这一思想．哥德巴赫猜想的命题是：

任何不小于6的偶数均可表成两个奇素数之和．

这个命题的证明遇到了极大的困难，两个多世纪以来一直无法证明．1900年，德国数学家希尔伯特在巴黎数学大会上提出了对20世纪的数学研究影响深远的23个数学难题，其中的第8个问题中就提到了哥德巴赫猜想．1912年德国数学家朗道在国际数学会的报告中曾悲观地说："即使要证明下面较弱的命题：'任何不小于6的自然数都能表示为不超过k（k为一个确定的整数)个素数之和'，也是现代数学力所不及的．"朗道的悲观论调并不正确，但他的话却提出了一种重要的数学思想：逐步逼近的方法．如果我们能找到一个这样的正整数k，不管它有多大，然后逐步缩小k的值，一旦达到$k=2$，问题即告解决．1950年尹文霖证明了$k\leqslant18$．

正在这条道路逐步逼近哥德巴赫猜想的同时，数学家们又从另一条道路逼近猜想．

假设自然数m的素因数的个数不超过a，就称m为素因数不超过a的殆素数．如果一个充分大的偶数可以表为两个素因数分别不超过a与b的殆素数之和时，则记为$(a+b)$．显然，如果能证明$(1+1)$，则哥德巴赫猜想即告证实．1957年，我国数学家王元证明了$(2+3)$．

上述结果虽然前进了许多，但人们还感到有点美中不足．即$(a+b)$中的两个数还没有一个是真正的素数．于是人们又转向从$(1+c)$(即充分大的偶数可以表示为1个素数和一个素因数不超过c的殆素数之和)向哥德巴赫猜想逼近．1966年，我国数学家陈景润证明了$(1+2)$．

在$(1+c)$这条战线上，我国数学家贡献最大，陈景润的

伟大成果离开摘取这颗“数学皇冠上的明珠”只有一步之遥了．

5．转化为已能解决的问题

这是转化思想中最重要也是最有效的思想之一．

我国当代数学家徐利治教授近年来致力于数学思想方法的研究，作出了杰出的贡献．他提出的“关系·映射·反演”方法，是把数学问题向已解决问题转化这一思想的典型范例．“关系·映射·反演”方法又称为 RMI 原理，它取关系(Relerionship)、映射(Mapping)、反演(Inversion)的第一个字母合成 RMI，故称 RMI 原理．

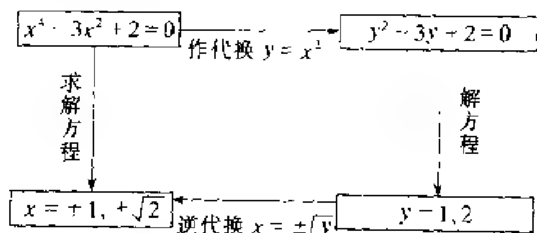
例如我们要解方程 $x^4-3x^2+2=0$，如果我们没有掌握四次方程的求解公式，但已熟悉二次方程的解法，便可以这样来解答：

(1) 作代换：$y=x^2$，得 $y^2-3y+2=0$．

(2) 解二次方程：$y_1=1$，$y_2=2$．

(3) 逆代换：$x=\pm1$，$\pm\sqrt{2}$．

这一过程可用框图表示如下：



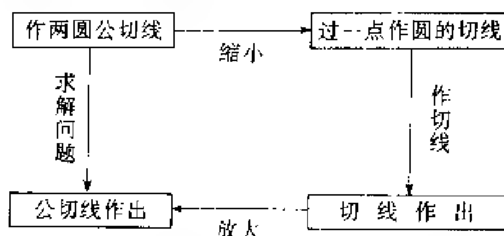
又例如，求作两个不等的圆的外公切线．通过圆外一点作圆的公切线是我们所熟知的，因此，可以把求作两圆的外公切线的问题转化为过圆外一点作圆的切线的问题．

(1) 缩小两圆的半径：将两圆的半径同时缩小，缩小的长度等于小圆的半径，则小圆化为一点．

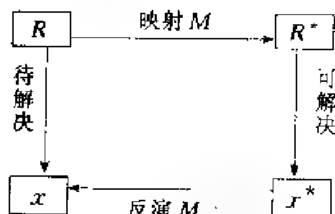
(2) 作特殊状态下的切线.

(3) 放大圆的半径，相当于将作出的切线平移一段距离，这段距离等于小圆的半径.

这个过程也可用框图表示如下：



把这种思想一般化就是 RMI 原理：令 R 表示由某些数学对象组成的关系结构系统，称为原象系统，其中有待确定的原象 x，但在 R 中直接确定 x 比较困难，但是可找到一种映射 M，通过 M 的作用把原象结构系统 R 映射为映象关系结构 R^*，在 R^* 中包含了原象 x 的映象 x^*，如果在 R^* 系统中有办法把 x^* 确定下来，然后通过反演即逆映射 M^{-1}，也就可把 x 确定下来. 用框图表示如下：



RMI 原理是一个普遍的思想方法，是一种最有效的转化.

例 人们很早就知道一元一次方程的求解方法，一元二次方程，化为典型状态之后，也有固定的解法. 至于一元三次方程的解法，却过了很久的时间才获得. 1545 年意大利数学家

卡丹在他出版的著作中公布了三次方程的求根公式，据称他是从意大利另一数学家塔塔烈处学来的.

卡丹的基本思想就是把三次方程转化为二次方程来求解，因为二次方程是已经解决了的. 怎样才能达到这一目的呢？他先从特殊状态考虑，对于特殊的三次方程

$$x^3+px+q=0, \qquad ①$$

如果令 $x=u+v$，则方程可化为

$$u^3+v^3+q+(3uv+p)(u+v)=0 .$$

如果再令 $3uv+p=0$，则得

$$\begin{cases} u^3+v^3=-q, \\ uv=-\dfrac{p}{3}; \end{cases} \quad 或 \quad \begin{cases} u^3+v^3=-q, \\ u^3v^3=-\dfrac{p^3}{27}. \end{cases}$$

因此，u^3 和 v^3 是二次方程

$$t^2+qt-\frac{p^3}{27}=0 \qquad ②$$

的两个根. 于是

$$u^3=-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}, \quad v^3=-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}},$$

从而 $x=u+v=\sqrt[3]{-\dfrac{q}{2}+\sqrt{\dfrac{q^2}{4}+\dfrac{p^3}{27}}}+\sqrt[3]{-\dfrac{q}{2}-\sqrt{\dfrac{q^2}{4}+\dfrac{p^3}{27}}}$.

既然对于特殊的三次方程①可转化为解典型的二次方程②，那么对于一般的三次方程

$$ax^3+bx^2+cx+d=0, \qquad ③$$

只要再设法把它转化为①就可以了，这只要作线性变换 $x=y-\dfrac{b}{3a}$，就可把方程③转化为方程①.

三、转化的途径

确定了转化的方向，还有一个如何实现转化的问题. 即通

过一条什么路线达到转化的要求，这当然要具体问题具体分析．现在将常见的转化途径列表如下．至于其具体的内容，将在第五、第六两章中详细论述．

（ ）直接转化(这一类转化的过程比较自然，转化的跨度比较紧凑、一般也不需要太多的构造过程，转化的方向也比较明确．总之，它比较“常规”．我们称它为第一类转化思想)

1．特殊化思想

(1) 特例思想

① 极端化思想

② 反例思想

③ 不变量思想

(2) 分类思想

(3) 分步思想

① 递进思想

② 递推思想

③ 逐步逼近思想——极限思想、离散化思想、逐步调整思想、逐步淘汰思想

2．一般化思想

(1) 模型化思想

(2) 强化命题思想

3．变换思想

(1) 数式变形和变量替换思想

(2) 几何变换思想

(3) 命题转化思想——分析、综合、间接证明

（二）横向转化(这一类转化的跨度较大，转化前后的状态之间缺乏必然的、紧密的联系．一般还需要构造某些转化的条

件．它是通常说的"非常规"问题，我们称它为第二类转化思想）．

1．映射思想

(1) 配对思想

① 倍数映射（包括一一映射）

② 对称映射

③ 构形映射

(2) 划分思想

① 染色

② 抽屉原理

③ 同余

(3) 赋值思想

① 布尔代数

② 特殊值法

③排序思想

(4) 表示思想

① 自然数的表示

② 坐标方法

③ 图表方法

(5) 同态思想

① 同态或同构

② 类比

2．数形结合思想

§3.3 构造思想概论

一、构造思想的意义及构造法的分类

构造思想是通过构造来建立数学理论、解决数学问题的一种数学思想．所谓构造，就是构建结构或体系，构造对象或指出达到某种目的的方式和途径．构造必须切实可行，它是直观的、定量的；并且必须能够在有限步骤内完成，即是能行的．

我们可以从不同的角度来考察、分析构造思想和构造法，即按不同的标准对构造进行分类．

1．构造对象与构造算法

根据构造的内容，构造法可分为构造对象和构造算法两种情形．

构造对象，就是把合乎某种要求的数学对象直接求出来，表达成明显的形式．

例如，为了证明"能够表示成一个平方数与一个质数之和的平方数有无穷多个"，我们可以构造无穷多个这样的数：$\left\{\frac{1}{4}(p+1)^2 \,\middle|\, p\text{ 为奇质数}\right\}$，由于质数有无穷多个，故我们构造的是一个无限集．又由

$$\left[\frac{1}{2}(p+1)\right]^2=\left[\frac{1}{2}(p-1)\right]^2+p,$$

故此集中的每个数都是平方数，且可以表成一个平方数与一个质数之和．

构造算法，就是指出求得某种合乎要求的对象或实现某种目的切实可行的有限的步骤．

例如，求最大公约数的"辗转相除法"、"开平方法"、几何作图中的"作法"等等，都是已给构造的算法．我国古算中的各种"术"(如"约分术"、"割圆术"、"方程术"…)也即各种各样的算法．

应该指出，"算法"一词不同于"计算"而有着十分广泛的涵义．首先，它并不单指计算，几何作图中的"作法"、博弈中的取胜"策略"、解线性方程组的"消去法"等等，都是算法的例子；其次，当要求的对象不能通过已知的对象明确地表示时，我们就以"算法"的形式指出求得未知对象的途径．例如，两个数的最大公约数不能通过这两个数本身来表示，而辗转相除法则告诉我们从这两数出发在有限步骤内可求得它们最大公约数的方法．线性规划中求线性约束下目标函数(也是线性函数)的极值的"单纯形法"、图论中求生成树、最小生成树、最短道路的算法，都属于此种情形．最后，在有些场合，构造算法与构造对象也没有绝对的界限．例如，一元二次方程的求根公式是一个代数式，因为一个代数式具有双重意义：既表示运算的过程，又表示运算的结果．从第一种涵义来理解，求根公式是一种算法：对一元二次方程的系数按公式所表明的顺序进行运算，就能求出方程的根；从第二种涵义来理解，求根公式作为一个运算结果整体，本身即为所构造的方程的根．

2．直接构造与递归构造

根据构造的方法，可分为直接构造与递归构造．

直接构造是一次完成，直接得到所要的结果．如前面关于平方数的例子即是一次构造了无穷多个合条件的平方数．

递归构造则是逐次完成、分步递进，其理论根据是数学归纳法原理．例如要证明对于任意的 n，存在 2^{n+1} 个由 1、2 组成的项数为 2^n 的有限序列，每两个序列至少有 2^{n-1} 个对应的

项互不相同．当 $n=1$ 时，$\{(1\ 1),\ (2\ 2),\ (1\ 2),\ (2\ 1)\}$ 是4个长度为2的序列，每两个至少有1个对应的项不同；若 A_n 是已构造的由 2^n 个由1、2组成的项数为 2^{n-1} 序列的集合，且 A_n 中的每两个序列至少有 2^{n-2} 个对应的项互不相同．对任意由1、2组成的有限序列 a，b，我们用 a' 表示将 a 中的1均换成2，2均换成1所得的序列，用 ab 表示由 a 和 b 互相连接所得的序列，且构造 A_{n+1} 为

$$A_{n+1}=\{aa,\ aa'\mid a\in A_n\}.$$

则 A_{n+1} 恰是 2^{n+1} 个由1、2组成的、项数为 2^n 的有限序列的集合，且 A_{n+1} 中的每两个序列至少有 2^{n-1} 个对应的项互不相同．由数学归纳法原理，我们对任意的自然数 n 完成了所需要的构造．由构造过程可以看出，递归构造的每一步构造，依赖于前面一步构造的完成．

另一种构造法表面上并不具备数学归纳法的形式，但本质上仍是逐次递进的，所以也可算为递归构造．例如问题"在各项都为自然数的等差数列中，或者没有完全平方数，或者有无穷多个完全平方数"可用构造法证明：若 $a_n=a_0+nd$ 是首项为 a_0、公差为 d 的等差数列中的一项，且 $a_n=b^2$（b 为自然数），则

$$(b+d)^2=b^2+(2b+d)d=a_0+(n+2b+d)d$$

是数列中的一项，这说明在数列中每个为完全平方数的项后面必然还有为完全平方数的项，故数列中含有无穷多个这样的项．

3．构造性定义和构造性证明

数学中常用构造法给定概念和确立命题，故有所谓概念的"构造性定义"和命题的"构造性证明"，它们反映了构造法在数学的不同方面的作用．

构造性定义指的是构造被定义的对象的方法，通常构造性定义有两种形式："生成定义"与"递归定义"，显然这种区分基于构造方法的不同。例如圆的生成定义是："将线段绕其固定的一个端点旋转一周时，另一个端点画出的曲线叫做圆。"图论中的"图"的定义也是构造性的：由非空有限集 V 及 V 中的元素的一些无序对组成的集合 E 构成的二元组 $G=(V, E)$ 称为一个(无向)图。递归定义则是数学归纳法用于定义概念，例如幂的概念可定义为

$$a^0=1;\ a^{n+1}=a\cdot a^n.$$

n 阶导数可定义为

$$\begin{cases} f^{(1)}(x)=f'(x)=\lim\limits_{\Delta x\to 0}\dfrac{f(x+\Delta x)-f(x)}{\Delta x}, \\ f^{(n+1)}(x)=[f^{(n)}(x)]'. \end{cases}$$

函数的差分可定义为

$$\Delta^{(1)}f(x)=\Delta f(x)=f(x+\Delta x)-f(x),$$
$$\Delta^{(n+1)}f(x)=\Delta(\Delta^{(n)}f(x)).\ (n\geqslant 1)$$

从这些例子可见，递归定义是一种递归构造。

构造性证明则是基于构造方法的一类证明，这类证明不满足于定性的结论，而要求是现实的、定量的。例如在证明某种对象(方程的根、合乎某种要求的多项式、等等)的存在性时，构造性证明要求将此对象具体构造出来。前面一些例子的证明就是构造性证明的例子。又如："证明：不能表示为一个质数与一个完全平方数之和的完全平方数有无穷多个"，亦可构造性地证明如后：考察集合 $\{n^2=(3k+2)^2\mid k=1, 2, \cdots\}$，这是一个无穷集合，若有

$$n^2=x^2+p,\ p\text{ 是质数}.$$

则由 p 是质数及

$$p = n^2 - x^2 = (n+x)(n-x)$$

得

$$n - x = 1,\ n + x = p.$$

故

$$\begin{aligned} p &= n + x = n + (n-1) = 2n - 1 \\ &= 2(3k+2) - 1 \\ &= 3(2k+1). \end{aligned}$$

这说明 p 不是质数，得出矛盾，故构造了无穷多个平方数不能表示为质数与平方数之和.

构造性证明是相对于非构造性证明而言的，两者的区别在于构造性证明实际上附带了能行性的要求. 例如，问题"证明存在两个无理数 x，y，使 $z = x^y$ 是有理数". 用非构造性方法可证明如下：反设对于任何两个无理数 x，y，$z = x^y$ 都是无理数，则$\sqrt{2}^{\sqrt{2}}$必为无理数. 于是$(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = (\sqrt{2})^2 = 2$ 是无理数，故得矛盾；用构造性的方法可证明如下：取 $x = \sqrt{2}$，$y = \log_2 9$，则 $x^y = (\sqrt{2})^{\log_2 9} = 2^{\log_2 3} = 3$，故合乎题意的无理数 x，y 存在. 比较两个证明可以看出，第一个证明实际上是构造反例证明原问题的否命题不成立，从而原问题的结论成立. 它并没有求得(构造出)使 $z = x^y$ 为有理数的无理数 x，y；第二个证明则直接构造出这样的无理数 x，y，使 $z = x^y$ 为有理数. 所以，用构造方法给出的证明，并不都是构造性证明，构造性证明的本质属性是能行性. 正是因为附加能行性的要求，构造性证明比非构造性证明含有更多的信息，因而更加受到重视！

4. 宏观构造与微观构造

根据构造的规模，可分为宏观构造与微观构造.

从《几何原本》开始，数学就是用逻辑演绎的方式表述的，

所以数学是一个逻辑演绎的系统。由于非欧几何的发现，人们认识到一个演绎系统必须有一定的基础、出发点或前提，并且可以选择不同的前提而形成不同的演绎体系。在这个基础上发展了数学的公理化方法，1899 年出版了德国数学家希尔伯特的名著《几何基础》。希尔伯特在这部著作中提出了“结合公理”、“顺序公理”、“合同公理”、“连续公理”、“平行公理”等五组共 20 条公理作为欧氏几何的基础，成功地建立了欧氏几何的公理系统。此外，皮亚诺关于自然数的公理系统，群、环、体等代数结构的公理系统，柯尔莫哥洛夫的概率论公理化结构等等相继出现。公理化方法在推动数学的发展中起了巨大的作用。

以布尔巴基为代表的现代结构主义学派发展了数学公理化的思想。公理化结构是一个数学系统(分支)的内部结构，结构主义者的理想则是要构造数学的整体结构。依照他们的观点，一个数学分支就是以一个非空集合为基础，在其中引入若干结构，然后展开理论探讨的。他们在众多的结构中抽出三种母结构：即代数结构、序结构、拓扑结构，一个集合可赋以这些母结构及其派生的结构中的一种或几种，而数学不同分支的差异在于其结构的差异。数学的发展，即各种结构的建成和发展。他们按照自己的理想“重建数学”，1939 年在法国巴黎出版了他们的著作《数学原本》(第一卷)，开始了这项伟大的重建工程，到 1973 年已出版 36 卷并且仍在继续出版，作为一个学派，他们的事业仍在发展之中。

无论是“公理化体系”还是“数学结构”，作为“体系”和“结构”的实现都是一种构造，这是宏观的构造，可见构造思想对于数学应该处在战略思想的位置。

至于微观的构造思想和构造方法，则贯穿于数学的定义、

证明、解题之中，几乎涉及数学的一切方面，我们已经举过不少例子，在下章中还将详细讨论。

数学的构造不同于机械的、物理的构造，它是一种纯粹思维性质的构造，存在于人的思维之中。构造的"材料"或"部件"是数学对象(如图形、概念、函数、公式、法则等等)，若干已知的对象通过逻辑手段"焊接铆合"，形成新的对象。整个构造的过程是一个思维过程，从"材料"到"产品"都极其抽象。正因为这样，数学中的构造富于技巧性，其精细、微妙是机械的构造所不能比拟的，其美学价值亦不逊于建筑、雕塑、绘画，并且品味更高。惟有懂得数学并具很高的修养方能欣赏，它反映了人类的智慧所达到的高度!

二、构造思想的发展和作用

1. 构造思想的发展和作用

构造思想对于数学是"与生俱来"，作为数学思想两大源泉的古代中国数学(以《九章算术》为代表)和古希腊数学(以《几何原本》为代表)，即有着明显的构造性特点和丰富的构造思想。中国的古算是由众多的"术"组成的算法系统，以至于古代称数学为"算学"；《几何原本》中也有求最大公约数、求勾股数等等许多算法，并且将作图题与证明题、计算题并列为几何的三类问题。

无论从宏观上看或微观上看，数学的发展都与构造不能分开。从宏观上看，一方面，由于数学的结构性的特点，每一种结构的构成或每一个体系的建立，都是一种宏观上的构造；另一方面，数学的发展，有时正是由于某种构造的发现所推动。古代的毕达哥拉斯学派认为"万物皆数"，他们所说的数是指整数，即宇宙间的一切现象都可以归结为整数或整数之比。按此

观点，一切线段皆可"通约"，这就是说，任意两线段 a、b（不妨设 $a<b$），以 a 量 b，量得 q_1 次而剩下一段 r_1：$0\leqslant r_1<a$，以 r_1 量 a，量得 q_2 次而剩下一段 r_2：$0\leqslant r_2<r_1$，如此继续，这个过程进行有限次必然结束，即必存在 n 使 $r_n=0$．这时 a 和 b 均为 r_n 的整数倍：$a=kr_n$，$b=hr_n$．容易看出，整个过程用算式表示，就相当于求两自然数的最大公约数的辗转相除法：

$$\begin{cases} b=q_1a+r_1, \\ a=q_2r_1+r_2, \\ \cdots \\ r_{n-1}=q_nr_n. \end{cases}$$

但是人们构造了不可通约的线段的例子：正方形的边与其对角线不可通约，否则$\sqrt{2}$就是一个有理数．这个例子的发现导致第一次数学危机．它告诫人们：直觉和经验并不可靠，必须重视演绎推理，几何学的演绎体系由此建立．至于微观构造，则在数学中比比皆是，构造法是每个数学研究工作者必须而且经常使用的一种手段．在下章中我们还将对微观构造作深层讨论．

由罗素悖论而引发的第三次数学危机，促使人们重视对于数学基础的研究，形成了逻辑主义、直觉主义、形式主义三大学派．因为矛盾出自集合论，逻辑主义、形式主义的代表人物罗素、希尔伯特等主张改造集合论，而直觉主义者则主张抛弃集合论，另搞一套．直觉主义强调心智活动中的构造性，认为"存在必须是被构造"，排斥传统数学中的非构造性部分．虽然他们把"构造方法"绝对化的做法并不妥当，但他们的工作推动了对于构造法的系统的研究．

构造思想和构造法也是联系数学理论及其应用的纽带．现代的自然科学和社会科学的许多部门都正在经历着一个"数学

化"的进程，马克思把数学化作为一门科学趋于成熟的标志，他说："一门科学只有当它达到了能够运用数学时，才算真正发展了。"但是要在一门科学以至任何一个实际问题中应用数学方法，都必须建立起适合这门科学或这个实际问题，又便于进行数学处理的数学模型，从这个意义上说，"数学化"就是"数学模型化"，而建立模型的过程就是一个构造的过程。

随着人类进入计算机时代，构造思想和构造性方法愈显示其重要性，因为计算机只能处理离散的数学模型，并且问题必须形式化和建立适当的算法，根据算法编制程序，计算机才能接受和运作，这就对数学的构造性方法提出了新的、更高的要求，组合学、离散数学、计算机科学、数值分析等构造性很强的数学分支得到很大的发展。

2．初等数学中常用的构造思想

因为本书主要讨论数学解题思想，重点在于微观构造的研究，下面我们将微观构造的常见内容列表如下，其详细的论证将在第四章中叙述。

（一）存在性的构造思想

1．直接构造

（1）几何作图

（2）其他构造

2．递归构造思想

（二）程序性的构造思想

1．程序构造思想

2．算法构造思想

（三）工具性的构造思想

1．构造图形

2．构造函数

3．构造辅助元素
4．构造辅助命题
5．构造模型
6．构造例与反例

第四章　构造思想

在这一章中，我们将较详细地讨论数学中的构造思想．我们从两个方面——数学本身的构造思想和数学解题中的构造思想展开分析．

§4.1　数学中的构造思想

构造思想是一种基本数学思想，它在数学中几乎到处都可以见到，因而构造思想在数学中的作用十分重要．

1．概念和理论的构造性

数学中的许多概念和理论，本身即是一种构造．

数是数学中最基本的概念．一方面，从人类实践的历史过程来看，人们是在数物体的个数（计数）中认识自然数；在分配物体中认识分数；从丈量长度中认识无理数；由运算的需要而引入虚数，进而认识复数与平面向量的关系，而使数的系统由自然数到有理数、到实数、到复数逐渐扩充的．另一方面，从数的概念的理论发展过程来看，数的概念的生成及每次的扩充都是一种构造．从集合论的观点来看，自然数是一列集合：

$$
\begin{aligned}
&0 = \Phi;\\
&1 = \{\Phi\};\\
&2 = \{\Phi,\{\Phi\}\};\\
&3 = \{\Phi,\{\Phi\},\{\Phi,\{\Phi\}\}\};\\
&\cdots
\end{aligned}
\qquad ①
$$

引入负整数后，自然数集 N 扩充为整数集 Z．这时，有理数可用整数对来构造：整数的有序对 $(a,b)(b\neq 0)$ 称为有理数，满足条件 $ad=bc$ 的两个有理数 (a,b)、(c,d) 视为是同一的（等价的、相等的）：$(a,b)=(c,d)$．有理数的运算定义为

$$
\begin{aligned}
(a,b)+(c,d)&=(ad+bc,bd),\\
(a,b)\times(c,d)&=(ac,bd).
\end{aligned}
\qquad ②
$$

将 (a,b) 写成 $\frac{a}{b}$，则它就是通常的分数，而②所定义的运算即通常的分数运算．当 $b=1$ 时，$\frac{a}{b}$ 就是整数 a，于是整数集 Z 扩充为有理数集 Q．由有理数集 Q 出发构造实数有几种相互平行的理论：实数可用有理数的基本序列（柯西序列）或有理数的分割（戴德金分割）来定义．复数可用实数的有序对来构造：称实数的有序对 (x,y) 为复数，复数的运算定义为：

$$
\begin{aligned}
(x_1,y_1)+(x_2,y_2)&=(x_1+x_2,y_1+y_2),\\
(x_1,y_1)\times(x_2,y_2)&=(x_1x_2-y_1y_2,x_1y_2+x_2y_1).
\end{aligned}
\qquad ③
$$

即 (x,y) 写成 $x+yi$，则它就是通常的复数，而③定义的运算即通常的复数运算，于是我们从自然数集 N 出发用构造的方法逐步扩充到复数集 C．由此可见，数的扩充的过程是一连串的构造的过程．

方程的研究在数学中占有很重要的地位．研究各类方程（代数方程、三角方程、指数方程、微分方程、积分方程、泛函方程…），首要的问题是方程的求解，即将方程的解直接构造出来．以整式方程为例，其解是由方程的系数决定的，而求解公式即是通过系数和运算构造方程的解的表达式．求出了这种表达式，则方程的性质也就一目了然．因此人们的兴趣首先在于寻找方程的求解公式．在 16 世纪已经得到了解三次方程

的卡丹公式和解四次方程的斐拉利公式，但人们在寻找五次及五次以上的高次方程的根式解中屡屡受挫．由于19世纪阿贝尔、伽罗华等人的工作，才证明了五次及五次以上的高次方程没有根式解．除求解公式而外，关于方程的研究工作还有另外的两个方面：其一，是关于方程的近似解法，即构造某种序列使其逐步逼近方程的解；其二，是关于方程的定性研究，即不必解方程，讨论方程的解的存在性、惟一性、解的分布等等．

2．问题性质和解答的构造性

数学中的许多问题具有构造的性质，这些问题须用构造的方法才能解决．

数学中由于数学理论的需要或者数学应用于实际问题的需要，常常要求寻求具有某种性质的数学对象（映射、函数、图形等等），只有将这些对象具体构造出来，问题才算解决．

例1 （线性回归）

设 X，Y 是两个相关的变量．为了从数学上刻画两个变量之间的关系，我们考察它们的对应值，设 $\{(x_i, y_i): i=1,2,\cdots,n\}$ 是一组对应值．如果我们已知 Y 是 X 的线性函数，或者通过在直角坐标系中描出各点 $(x_i, y_i)(i=1,2,\cdots,n)$，从得到的"散点图"中发现这些点大致分布在某一直线的两侧，从而认定在所考察的范围内 Y 与 X 有近似的线性关系，则我们可用

$$Y=aX+b$$

来表示这种关系，其中 a，b 为待定参数．虽然两点决定一条直线，但由于观察值总有误差，故我们从 n 个点中取出两点以确定参数 a，b 是不恰当的；应当充分地利用观察所得的所有信息，使我们得到的直线在总体上与所有的点的"接近的程度"最好，这条直线称为"回归直线"．回归直线上横坐标为 x_i

的点是(x_i, y_i')，其中

$$y_i' = ax_i + b.$$

量$(y_i - y_i')^2$反映了在$x = x_i$时观察点(x_i, y_i)与回归直线上的对应点(x_i, y_i')之间的差别，设

$$Q = \sum_{i=1}^{n}(y_i - y_i')^2 = \sum_{i=1}^{n}[y_i - (ax_i + b)]^2,$$

则Q的大小从总体上刻画了回归直线与观察点$\{(x_i, y_i): i = 1,2,\cdots,n\}$的接近程度，回归直线使这种接近程度最好，故应使$Q$取极小值．我们可根据这一要求来构造回归直线．从而求得

$$\begin{cases} b = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \Big/ \sum_{i=1}^{n}(x_i - \bar{x})^2, \\ a = \bar{y} - b\bar{x}. \end{cases}$$

其中

$$\begin{cases} \bar{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i, \\ \bar{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i. \end{cases}$$

所以线性回归问题就是构造回归直线，即求出回归直线的方程．

例 2 （插值公式）

多项式是最简单的函数，一般的函数可以用多项式进行拟合．设$y = f(x)$是任意函数．对于自变量的一组值$\{x_i: i = 0, 1, 2, \cdots, n\}$，已知函数的对应值是$\{y_i = f(x_i), i = 0, 1, 2, \cdots, n\}$，则$\{P_i = (x_i, y_i): i = 0, 1, 2, \cdots, n\}$是函数$y = f(x)$的图象上的一列点．插值问题的提法是：求$n$次多项式$P(x)$，使

$$P(x_i) = f(x_i), \quad i = 0, 1, 2, \cdots, n$$

即使 $y=P(x)$ 的图象与已知函数 $y=f(x)$ 的图象在点 P_i（$i=0,1,\cdots,n$）处重合．所求的这个多项式可以用迭加的方法构造：首先，对每个 i 构造多项式 $P_i(x)$，使

$$P_i(x_i)=f(x_i),P_i(x_j)=0\ (j=0,1,\cdots,n,\ j\neq i).\ (*)$$

则

$$P(x)=\sum_{i=0}^{n}P_i(x)$$

即为所求．取

$$P_i(x)=f(x_i)\frac{(x-x_0)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_0)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)},$$

则 $P_i(x)$ 适合（*），故求得

$$P(x)=$$

$$\sum_{i=0}^{n}f(x_i)\frac{(x-x_0)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_0)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)}.$$

构造所得的多项式 $P(x)$ 称为拉格朗日内插公式．

3．存在性与构造性

构造的方法是解决存在性问题的直接方法．

在定义一个数学概念或一种数学结构以后，首要的问题是这个概念的外延(即概念所概括的对象)及这种数学结构的存在性．存在性是赖以进行讨论的前提，不首先解决存在性问题，则理论的构建缺乏基础，一切讨论都将落空．解决存在性问题的直截了当的方法，就是将要证明存在的对象直接构造出来．

例 1 （切比晓夫多项式）

在数值分析中，我们常将次数较高的多项式用次数较低的多项式代替以简化运算．设 n 次多项式 $P(x)$ 的首项系数是 $a_0\neq0$，取一个首项系数为 1 的 n 次多项式 $T_n(x)$，则 $P_1(x)=P(x)-a_0T_n(x)$ 的次数低于 n．这个过程称为多项式的精简，

精简后获得的多项式$P_1(x)$的次数比$P(x)$低，但我们还希望在一定的范围(例如$(-1,1)$)内$P_1(x)$与$P(x)$相差最小．显然，当$T_n(x)$在$(-1,1)$上与0偏离最小时，从总体上看$P_1(x)$在$(-1,1)$上与$P(x)$最接近．

寻找在$(-1,1)$上与0偏离最小的多项式的问题是上世纪后半期伟大的俄罗斯数学家切比晓夫提出的．这样的多项式就称为切比晓夫多项式．

切比晓夫多项式的严格的数学定义是：设全体首项系数为1的n次实系数多项式组成的集合为P_n．对任意的$f(x)\in P_n$，记

$$m(f)=\max_{-1\le x\le 1}|f(x)|,$$

则$m(f)$表示在闭区间$[-1,1]$上多项式$f(x)$的值与0的最大偏离．P_n中使$m(f)$达到极小值的多项式称为n次切比晓夫多项式，通常用$T_n(x)$表示．

很明显，切比晓夫多项式的定义本身并不表明这种多项式的存在．要使这个定义有意义，必须证明存在性；要使这样定义的切比晓夫多项式真正用于数值分析中多项式的精简，还必须将多项式$T_n(x)$具体构造出来．

下面我们扼要地介绍这个问题的一种解法．它是莫斯科大学的姆洛杰耶夫教授给出的．

要构造切比晓夫多项式，首先要认识这个多项式应具备的性质，也就是说，$T_n(x)$如果存在，则应该满足一些什么必要条件．研究表明，$T_n(x)$应具有下面的性质：

性质1　$T_n(x)$在$[-1,1]$上恰有n个相异实根．

性质2　$|T_n(x)|$在$[-1,1]$上的极大值均相等．

进而可以证明，性质1、性质2也是决定切比晓夫多项式

的充分条件(因而是充要条件)．于是可以依据这两个性质来构造切比晓夫多项式．容易验证，

$$f_n(x)=\frac{1}{2^n}\cos(n\arccos x)$$

恰好具备上述两个性质，同时又可证明切比晓夫多项式如果存在必定惟一，于是$f_n(x)$就是我们所要构造的$T_n(x)$：$T_n(x)=f_n(x)$．通过适当的变形以消去$f_n(x)$表达式中的余弦和反余弦符号，可以将$f_n(x)$还原成多项式的形式而求得

$$T_n(x)=\frac{1}{2^{n-1}}[x^n-C_n^2x^{n-2}(1-x^2)+C_n^4x^{n-4}(1-x^2)^2+\cdots].$$

至此，n次切比晓夫多项式构造成功，因而其存在性问题得以解决．

例2 (罗氏几何的相容性)

罗巴切夫斯基用“罗氏公设”代替欧几里得的“第五公设”(即平行公理)，从而创立了“罗氏几何”．“罗氏公设”表述为：“过平面上一已知直线外的一点至少可以引两条直线与该已知直线不相交.”这一公设显然与直觉相悖，因而不能不引起对这个公理系统的存在性(合理性)的怀疑．虽然一时尚未发现公理系统中的矛盾，但谁也无法担保将来不会出现矛盾．于是数学家们致力于罗氏公理系统的相容性(无矛盾性)的证明，相容性是公理系统能够合理存在的先决条件．

但是如何证明公理系统的相容性确是一大难题．大家认为，既然没有人对欧氏几何表示怀疑，则欧氏几何的相容性是大家公认的．如果我们在欧氏几何中构造一组对象，将其解释为罗氏几何中的“点”、“直线”等等，并且对对象之间的关系(如“相交”、“平行”等等)作出具体界定，然后验证按照所作的解释和界定，罗氏几何的所有公理均得以满足，则罗氏几何的

相容性归结为欧氏几何的相容性．因为罗氏几何如果出现矛盾，则将导致欧氏几何出现矛盾．概而言之，只要能在欧氏几何中构造出罗氏几何的模型，则可以确认罗氏几何对于欧氏几何的相对相容性．

历史上曾出现过多种罗氏几何模型，如贝特拉米模型、克莱因模型、庞加莱模型．此处仅简单介绍克莱因模型．考察欧氏几何平面上的一个开圆(即一个圆的内部而不含圆周)．开圆内的点作为罗氏几何的点，开圆的弦(无端点)解释为罗氏几何的直线．整个开圆即罗氏平面．而罗氏公设在模型中对应为"过圆内不在已知弦上的点，至少可以引两条弦不与已知弦相交．"在图4－1中，PA和PB过P而不与已知弦相交(注意A，B不是罗氏几何的点)，且PA、PB把过P点的罗氏几何直线分为两类，一类与已知直线(即弦AB)相交，另一类不相交．而$\angle APB$即平行角．因而罗氏公设在克莱因的模型中成立．罗氏公理系统的其余公理都与欧氏几何相同．在欧氏平面上构造的克莱因模型自然满足这些公理．因而克莱因模型是罗氏几何的一个十分简洁的模型，构造这个模型使我们从理性上认识到罗氏几何存在的合理性．

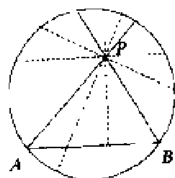


图4－1

4．解题方法的构造性

构造性方法是相对于传统数学中的非构造性方法而言的．

非构造性方法只涉及对象的存在性，目标的可实现性，因而是一种定性的方法．例如"任何多项式可惟一地分解为不可约多项式之积"、"任何非常数的复系数多项式至少有一个复数根"(代数基本定理)，"连续函数在闭区间上必有最大值和最小值"…都是通过非构造性方法给出证明的．非构造性方法只涉

及可分解性、根与最值的存在性，而不需要因式分解的实际步骤和求根、求最值的具体方法．对于代数基本定理，非构造性方法只是证明了"无零点的复系数多项式必为常数"，并不直接证明非常数的复系数多项式有根，更不曾把这个存在的根实际找出来或指明一种求根的方法，便已经断言定理成立．

构造性方法则不满足于对象的存在性和目标的可实现性的结论，而进一步要求明确指出构造对象的具体方法和实现目标的具体步骤，这种方法和步骤还必须具有能行性，即可以在有限步骤内实际完成．例如，克朗内克具体构造了一种实际方法，按照他的方法实施，任何有理系数多项式的因式分解可以在有限步内实现．布劳威尔则用构造法证明了代数基本定理，他的方法确能求出任意给定的非常数复系数多项式的根．

传统数学中也有构造性方法，例如，平面几何中有垂线、角平分线等概念，也有作垂线、作角平分线的作图方法；二次方程的求根公式可以实际求出任意二次方程的根；在定义两个自然数的最大公约数的概念之后，有求最大公约数的欧几里得算法(辗转相除法)．我国古代数学在构造法研究方面也卓具成就．但直觉主义者出于重建数学基础的需要，把构造法推向极端，按照他们的观点，"存在必须是被构造"．他们对每个数学命题都附加上"能行性"的要求，向传统数学提出挑战．他们的工作，一方面大大地丰富了传统数学的内容，开拓了数学研究的新领域，但另一方面，他们又否认了一大批有重要意义的数学成果，未免有失偏颇．事实证明：构造性方法与非构造性方法兼容并存的情况，有利于数学的发展．

例 1 （克朗内克关于有理系数多项式因式分解的算法）

任何有理系数多项式总可以表示成一个常数(有理数)与一个整系数多项式之积，故我们只需讨论整系数多项式的因式

分解．

设$f(x)$是一个n次整系数多项式．若$f(x)$可以分解，那么必有一个因式的次数不超过$s=\left[\frac{n}{2}\right]$（$[a]$表示不超过$a$的最大整数），故我们需要考察$f(x)$是否具有次数不超过$s$的因式．

任意取定$(s+1)$个整数值$a_0,a_1,\cdots,a_s$，求出多项式的值$f(a_0),f(a_1),\cdots,f(a_s)$，它们都是整数．若$f(x)$有因式$g(x)$，则$f(a_0)$必被$g(a_0)$整除，$f(a_1)$必被$g(a_1)$整除，$\cdots$，$f(a_s)$必被$g(a_s)$整除．但每个整数$f(a_i)(i=0,1,\cdots,s)$可分解质因数：

$$f(a_i)=\pm p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_r^{\alpha_r},$$

$$(0<p_1<p_2<\cdots<p_r\ \text{均为质数})$$

从而$f(a_i)$的因数有$2^n(\alpha_1+1)(\alpha_2+1)\cdots(\alpha_r+1)$个．故每个$g(a_i)$只有有限多种可能的取值，因而$g(a_0),g(a_1),\cdots,g(a_s)$的值也只有有限多种可能的组合．

对每组值$g(a_0),g(a_1),\cdots,g(a_s)$，利用拉格朗日插值公式（§4.1 例2）可以构造惟一的一个多项式，这样可以确定有限多个$g(x)$，它们是$f(x)$的次数不超过s的可能因式．于是我们就求出了$f(x)$的次数不超过s的可能的因式，它们只有有限多个，但已包括了$f(x)$的所有次数不超过s的因式．

现在应用除法算式，检验每个$g(x)$是否真是$f(x)$的一个因式．若每个$g(x)$均不能整除$f(x)$，则$f(x)$是素多项式，因而不可分解；若某个$g(x)$整除$f(x)$，则已经找到了$f(x)$的一个因式$g(x)$，于是有$f(x)=g(x)f_1(x)$，此时$g(x)$、$f_1(x)$的次数均低于$f(x)$的次数．

对$g(x)$、$f_1(x)$同时重复上面的手续，则$f(x)\ \ g(x)$．

$f_1(x)$可继续分解．由于$f(x)$的次数有限，故这样的手续不可能无限地重复下去，即$f(x)$的因式分解必在有限步内完成．

例 2 （关于图形的组成相等）

如果用一定的方式把两个图形中的一个剖分成有限个部分以后，又可以用另外的方式配置这些部分，以组成这两个图形中的另一个，那么这两个图形叫做"组成相等"．

显然，组成相等的图形必有相等的面积．有趣的是：两个面积相等的多边形必也组成相等．这一事实的证明由以下几个引理组成．注意到每个引理的证明都是构造性的，所以证明中实际上已给出剖分两个面积相等的多边形中的一个以组成另一个的具体方法．

引理 1　若图形 A、B 都与图形 C 组成相等，则 A 与 B 组成相等．

证明：如图 4－2 所示．设按实线剖分 C 所得的各部分可组成A；按虚线剖分C所得的各部分可组成B，则同时用实线和虚线将 C 剖分成更细的部分，它们既可以组成 A，又可以组成 B．由此显然可知 A，B 组成相等．

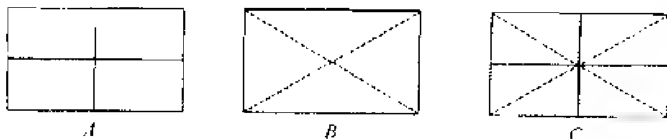


图 4－2

引理 2　任何三角形与一个矩形组成相等．

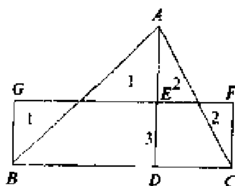


图 4－3

证明：设角 A 是△ABC 的最大内角，则高 AD 必在△ABC 的内部．E 是

AD 的中点，过 E 作 $GF \mathrel{\underline{\parallel}} BC$，使其构成矩形 $BCFG$ ．则将三角形剖分成两个三角形 1 和 2 及一个梯形 3 后，由 1，2，3 可组成矩形 $BCFG$（如图 4－3）．

引理 3　共有相同底边且面积相等的两个平行四边形组成相等．

证明：设平行四边形 $ABCD$ 与 $ABEF$ 有公共边 AB ．因为它们面积相等，故另一对边 CD、EF 在与 AB 平行的同一条直线上．如图 4－4 所示，在两条平行线之间作两组两两相邻的平行四边形，分别与已知平行四边形 $ABCD$、$ABEF$ 全

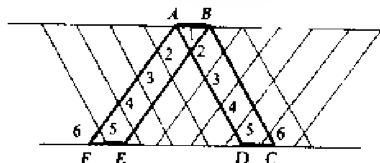


图 4－4

等．则已知平行四边形各边被剖分成若干部分．最上面的一部分是共有的一个三角形，各部分从上到下依次编号．将 $ABCD$ 逐步左移一步，两步，三步，…，每步之长等于线段 AB，则第二部分、第三部分、第四部分…将依次重合，于是可知两个平行四边形组成相等．

引理 4　面积相等的两个矩形组成相等．

证明：设矩形 $ABCD$、$EFGH$ 面积相等．设 AB 是 AB，

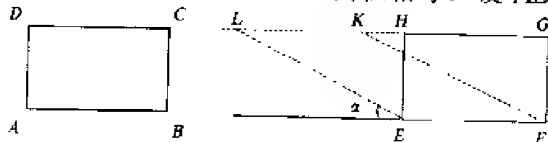


图 4－5

BC、EF，FG 中最长的一条线段．以 EF 为一边作平行四边形$EFKL$，使 $LE = AB$ 且与 $ABCD$、$EFGH$ 有相同的面积．$EFGH$、$EFKL$ 共有底边 EF 且面积相等，故由引理 3，它们组成相等；$ABCD$、$EFKL$ 有相等的边 $AB = LE$，故由引理 3，它们也组成相等．$ABCD$、$EFGH$ 均与 $EFKL$ 组成相等，故由引理 1，它们组成相等．

引理 5　任何多边形与一个矩形组成相等．

证明：将多边形剖分成若干三角形．用 1，2，…，n 编

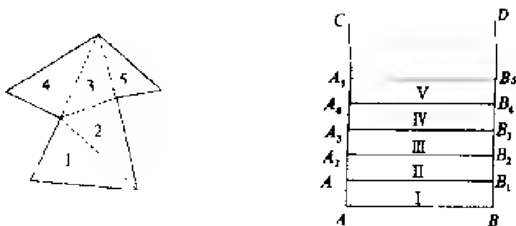


图 4－6

号．作矩形 ABB_1A_1 与编号为 1 的三角形面积相等；再在 ABB_1A_1 的外部作矩形 $A_1B_1B_2A_2$ 与编号为 2 的三角形面积相等，…，作矩形 $A_{n-1}B_{n-1}B_nA_n$ 与编号为 n 的三角形面积相等，得矩形 ABB_nA_n．

编号为 i 的三角形与一个矩形组成相等(引理 2)，这个矩形与编号为 i 的三角形面积相等，因而与矩形 $A_{i-1}B_{i-1}B_iA_i$ 面积相等，故这两个矩形亦组成相等(引理 4)．于是编号为 i 的三角形与矩形 $A_{i-1}B_{i-1}B_iA_i$ 组成相等($i = 1, 2, \cdots, n$)．由此可知原多边形与矩形 ABB_nA_n 组成相等．

现在证明两个面积相等的多边形必组成相等．由引理 5，每个多边形都与一个矩形组成相等．由多边形面积相等可知两个矩形也面积相等，于是这两个矩形组成相等(引理 4)．故由

引理1，这两个多边形组成相等.

5. 用构造的观点分析数学对象的结构

从复杂事物与简单事物的联系中认识复杂事物，从未知事物与已知事物的联系中认识未知事物，这是认识的一条基本途径．用构造的观点分析数学对象的结构，有助于把握对象的实质和不同对象之间的联系．

数学中有许多所谓分解定理．例如，自然数可以惟一地分解为质数的乘积(算术基本定理)；多项式可以惟一地分解为不可约多项式的乘积；任何向量可以惟一地分解为单位坐标向量的线性组合；任意函数$f(x)$可以分解为奇函数$\frac{1}{2}[f(x)-f(-x)]$与偶函数$\frac{1}{2}[f(x)+f(-x)]$之和．这些分解定理揭示了数学对象的内部结构和表现形式．

例1 （广义等比数列的结构）

称数列$\{x_n\}$为广义等差数列，如果对任意的$n\geqslant1$，均有

$$x_{n-1}+x_{n+1}=ax_n\ (a\ 为常数). \quad ①$$

称数列$\{x_n\}$为广义等比数列，如果对任意的$n\geqslant1$，均有

$$x_{n-1}x_{n+1}-x_n^2=M\ (M\ 为常数). \quad ②$$

显然，当$a=2$，$M=0$时，广义等差数列、广义等比数列就是通常的等差数列、等比数列．容易看出，①是一个线性关系式，而②是二次式．故从定义看，广义等差数列较广义等比数列简单一些．事实上，广义等差数列是一类二阶线性递归数列，其递推方程为

$$x_{n+1}=ax_n-x_{n-1}\quad (n\geqslant1).$$

有趣的是，广义等差数列都是广义等比数列．事实上，设$\{x_n\}$是广义等差数列，则由①可得

$$x_{n-1}x_{n+1}-x_n^2$$

$$= x_{n-1}(ax_n - x_{n-1}) - x_n^2$$
$$= (ax_{n-1} - x_n)x_n - x_{n-1}^2$$
$$= x_{n-2}x_n - x_{n-1}^2.$$
于是可知此式的值不随 n 的改变而改变，即

$$x_{n-1}x_{n+1} - x_n^2 = M \text{（常数）}.$$

故 $\{x_n\}$ 是广义等比数列。

反之，如果广义等比数列也都是广义等差数列，则两者具有相同的外延，因而是同一个概念，然而事实并非如此！我们只能证明各项均不为 0 的广义等比数列是广义等差数列。事实上，由②可得

$$x_{n-1}x_{n+1} - x_n^2 = x_{n-2}x_n - x_{n-1}^2,$$

故 $$x_{n-1}(x_{n-1} + x_{n+1}) = x_n(x_{n-2} + x_n).$$

但 x_{n-1}，x_n 均不为 0，故

$$\frac{x_{n-1} + x_{n+1}}{x_n} = \frac{x_{n-2} + x_n}{x_{n-1}}.$$

于是对任意的 n

$$\frac{x_{n-1} + x_{n+1}}{x_n} = a \text{（常数）},$$

即对任意的 n，均有

$$x_{n-1} + x_{n+1} = ax_n.$$

故 $\{x_n\}$ 是广义等差数列。

这一事实的重要意义在于揭示了广义等比数列的结构：广义等比数列由若干各项均非 0 且满足不同递归关系的有限广义等差数列相互连接而成，在相邻的两个有限广义等差数列之间恰有一项为 0；如果广义等比数列出现相邻的两项同时为 0，则由②可知，从这两项开始以后的所有的项全部为 0 。

例 2 （模 m 的斐波那契数列的结构）

适合

$$\begin{cases} F_{n+2}=F_{n+1}+F_n, & (n\geqslant 0) \\ F_0=0,\ F_1=1 \end{cases} \qquad ①$$

的二阶递归数列 $F=\{F_n\}_{n\geqslant 0}$ 称为斐波那契数列．设 $m\geqslant 2$ 为自然数，令

$$d(m)=\min\{n: m\mid F_n\}$$

即 $F_{d(m)}$ 是数列 F 中第一个被 m 整除的项．则可以证明 $\{F_n\}$ 具有性质

$$m\mid F_n \Leftrightarrow d(m)\mid n \ . \qquad ②$$

以 m 除 F_n 并取最小非负剩余 $F_n(m)$ $(0\leqslant F_n(m)<m)$，则称数列 $\{F_n(m)\}$ 为模 m 的斐波那契数列．显然，$\{F_n(m)\}$ 关于模 m 满足递推关系①．

今考察数列 $\{F_n(m)\}$ 的结构．

数列 $\{F_n(m)\}$ 的一个重要特性是周期性．设 $L(m)$ 是这个数列的最小正周期．则 $F_{L(m)}(m)=F_0(m)=0$，即 $m\mid F_{L(m)}$，由②知 $d(m)\mid L(m)$．记 $r=L(m)/d(m)$．

设 $t=F_{d(m)-1}(m)$，则由 $F_{d(m)}(m)=0$ 及递归关系①，可知

$$F_{d(m)+1}(m)=F_{d(m)}(m)+F_{d(m)-1}(m)=t,$$

$$F_{d(m)+2}(m)=F_{d(m)+1}(m)+F_{d(m)}=t \ .$$

但 $F_1(m)=F_2(m)=1$，故知

$$F_{d(m)+1}(m)=tF_1(m),\ F_{d(m)+2}(m)=tF_2(m).$$

于是由递归关系①的齐次性可知，对任意 k，

$$F_{d(m)+k}(m)\equiv tF_k(m) \pmod m.$$

因而对任意非负整数 j，

$$F_{jd(m)+k}(m)=t^jF_k(m),$$

依次让 $j=0, 1, \cdots, r-1$，并对每个固定的 j $(1\leqslant j\leqslant r-1)$，

让 $k=0, 1, \cdots, d(m)-1$, 我们得到$\{F_n(m)\}$的一个周期. 将其排列为

$$\begin{cases} 0,1,1,F_3(m),\cdots,F_{d(m)-1}(m)\equiv t, \\ 0,t,t,tF_3(m),\cdots,tF_{d(m)-1}(m)\equiv t^2, \quad (\bmod\ m) \\ \cdots \\ 0,t^{r-1},t^{r-1},t^{r-1}F_3(m),\cdots,t^{r-1}F_{d(m)-1}\equiv t^r\equiv 1. \end{cases}$$

容易看出,这样排列的每一列是一个等比数列,即$\{F_n(m)\}$的子序列$\{F_{jd(m)+k}(m)\}$ $(0\leqslant k\leqslant d(m)-1)$, 是以$F_k(m)$为首项, t 为公比的等比数列:

$$F_{jd(m)+k}(m)\equiv F_k(m)t^{j-1}(\bmod\ m),$$

因而整个$\{F_n(m)\}$是由分别以$F_0(m)$, $F_1(m)$, $\cdots$, $F_{d(m)-1}(m)$为首项并且有相同公比 t 的$d(m)$个模的等比数列的项依次交替排列而成的.

上面两例讨论的是一类对象的个体结构, 用构造的观点讨论一类对象的整体结构同样也是十分有趣而且是有意义的.

例 3 (二阶等差数列的结构)

设$\{x_n\}$是一个序列, $\{n_i\}$是自然数列的一个子列, 且 $n_0=0$, 则称序列 $\left\{\sum\limits_{k=n_i+1}^{n_{i+1}} x_k\right\}_{i\geqslant 0}$ 为序列$\{x_n\}$关于序列$\{n_i\}$的一个划分. 特别, 当$\{n_i\}$为等差数列时的划分称为数列的等差划分.

可以证明, 当$\{x_n\}$是等差数列时, $\{x_n\}$的等差划分是二阶等差数列. 这一事实可概括地说成"等差数列的等差划分是二阶等差数列". 有人猜想:"二阶等差数列都是等差数列的等差划分", 但很快发现这一猜想不能成立, 于是要问:"什么样的二阶等差数列是等差数列的等差划分?" 有人给出了二阶等

差数列是等差数列的等差划分的充要条件．三阶等差数列可用其通项公式来刻画：一个数列是三阶等差数列，当且仅当其通项公式$f(n)$是n的三次多项式．因此上面的充要条件应当加在$f(n)$的系数上，其结论可叙述为："已给三阶等差数列$\{f(n)\}$：

$$f(n)=An^3+Bn^2+Cn+D \quad (A\neq 0),$$

则$\{f(n)\}$为等差数列的等差划分，当且仅当下面的两条同时成立：

i) $H \triangleq B/3A$ 是大于 -1 的有理数；

ii) $G \triangleq 2B^3-9ABC+27A^2D=0$．

由此可知，在全体三阶等差数列中，只有一部分是等差数列的等差划分．从理论的完整性来考虑，我们自然要问：另一部分三阶等差数列，即不成为等差数列的等差划分的三阶等差数列，与等差数列的等差划分有什么关系？回答了这个问题，就得到了全部三阶等差数列的一个完整的刻画．

这个结果，实际上可以总结为下面的三条结论：

1) 当 H 为大于 -1 的有理数时，

若 $G=0$，则$\{f(n)\}$是等差数列的等差划分；

若 $G\neq 0$，则$\{f(n)\}$是等差数列的等差划分与非零常数列$\{G/27A^2\}$的和．

2) 当 H 为不大于 -1 的有理数时，若不计数列的前若干项，则

当 $G=0$ 时，$\{f(n)\}$从某项开始以后的部分是等差数列的等差划分；

当 $G\neq 0$ 时，$\{f(n)\}$从某项开始以后的部分是等差数列的等差划分与非零常数列$\{G/27A^2\}$的和．

3) 当 H 为无理数时，$\{f(n)\}$为等差数列的等差划分与一

个二阶等差数列的和。

这三条结论深刻地揭示了全体三阶等差数列组成的集合的结构，故可以称为"三阶等差数列的结构定理"。其内容可以很清晰地概括在下表之中。

关于三阶等差数列的结构定理

G \ H	H 为大于 1的有理数	H 为不大于 −1的有理数	H 为无理数
G = 0	F 为等差数列的等差划分	划去 F 的前若干项余下的数列是等差数列的等差划分	F 为等差数列的等差划分与二阶等差数列的和
G≠0	F 为等差数列的等差划分与非零常数列{G/27A²}的和	划去 F 的前若干项，余下的数列是等差数列的等差划分与非零常数列{G/27A²}的和	

$F = \{f(n)\}: f(n) = An^3 + Bn^2 + Cn + D$,

$H = B/3A, G = 2B^3 - 9ABC + 27A^2D$.

§4.2 数学解题中的构造思想

构造思想在数学中占有十分重要的地位，在上节中我们已对此进行了多方面的分析。同样地，构造思想在数学解题中亦有着十分重要的作用，这是本节我们所要讨论的问题。

构造思想在数学解题中的作用主要表现在两个方面：其一，许多数学问题本身具有构造性的要求，或者可以通过构造而直接得解；其二，许多问题，如果通过构造相应的数学对象（如映射、函数、数列、图形、命题等等）作为辅助工具，则问

题容易获得解决.

一、用构造方法解数学问题

1. 构造性问题

这类问题的目标，就是要求构造适合某些条件的数学对象. 所以这类问题的求解就是将要求的对象直接构造出来，然后证明它合乎要求的条件.

例 1 求作集合 A_1, A_2, …, A_n, 使其中的任意 k 个集合的交不空，任意$(k+1)$个集合的交是空集，并使并集 $A=\bigcup_{i=1}^{n}A_i$ 中所含的元素最少.

解：设 $N=\{1,2,\cdots,n\}$. N 中的 k 个相异元素的无序组 $(a_1,a_2,\cdots,a_n)$称为 N 的一个 k 元组合. N 的全体 k 元组合(共有 C_n^k 个)组成集合 A:

$$A=\{(a_1,\cdots,a_k)|a_i\in N,i=1,2,\cdots,k\} \qquad ①$$

对每个 i $(1\leqslant i\leqslant n)$, 令 A_i 是 N 的含数 i 的全体 k 元组合组成的集合：

$$A_i=\{(a_1,\cdots,a_k)\in A\,|\,i\in(a_1,\cdots,a_k)\} \qquad ②$$

则 A_1, A_2, …, A_n 即为所求.

事实上，显然 $A=\bigcup_{i=1}^{n}A_i$, 且对任意的 $1\leqslant r_1<r_2<\cdots<r_k\leqslant n$, $\bigcap_{i=1}^{k}A_{r_i}=\{(r_1,r_2,\cdots,r_k)\}\neq\varnothing$; 而对任意的 $1\leqslant r_1<r_2<\cdots<r_k<r_{k+1}\leqslant n$, 因为每个 k 元组合不可能同时含有$(k+1)$个相异元素 r_1, r_2, …, r_k, r_{k+1}, 故 $\bigcap_{i=1}^{k+1}A_{r_i}=\varnothing$. 又若集合 A_1', A_2', …, A_n'也具有所要求的性质，即

i) A_1', …, A_n'中任意k 个的交不空；

ii) A_1', …, A_n'中任意$(k+1)$个的交为空集，

则对于 A_1'，在 $(n-1)$ 个集合 A_2'，…，A_n' 中任取 $(k-1)$ 个与 A_1' 的交不空．于此交中取定一元，可得 A_1' 的 C_{n-1}^{k-1} 个元，由 ii)，这些元素互不相同；用同样的方法可在每个 $A_i'$$(i=2,\cdots,n)$ 中取得 C_{n-1}^{k-1} 个不同的元，故我们一共得到 nC_{n-1}^{k-1} 个元，这些元素中当然可能有些会重复出现，但由 ii)，每个元素重复出现的次数不会超过 k 次，因而其中相异的元素不会少于 $nC_{n-1}^{k-1}/k=C_n^k$ 个．故知

$$\bigcup_{i=1}^{n} A_i' \text{中的元素个数} \geqslant C_n^k.$$

综上述；我们所构造的 A_1，…，A_n 满足问题的要求．

例 2 由 0 和 1 组成的有序组作元素所构成的集合称为基本集，如果任何一个由 0 和 1 组成的无穷序列中都存在由若干相邻项组成的有序组属于这个集合，求作由 51 个含 8 个数字(0 或 1)的有序组构成的基本集．

解：作集合

$$A=\{a_1,a_2,a_3;b_i(i=1,\cdots,32),c_j(j=1,\cdots,16)\},$$

其中 $a_1=00000000$，$a_2=11111111$，$a_3=01010101$；

$b_i=100\beta_i$，β_i 为由 5 个数字(0 或 1)组成的有序组；

$c_j=1101r_j$，r_j 为由 4 个数字(0 或 1)组成的有序组，

则 A 含有 $3+32+16=51$ 个元素，且每个元素都是由 8 个数字(0 或 1)组成的有序组．下面证明 A 是基本集．

设 $x=(x_1,x_2,\cdots,x_n,\cdots)$ 是由 0，1 组成的无穷序列，则有以下三种可能情形：

1) x 从某处开始全为 0：这时 x 中含有 a_1 作为其一段；

2) x 从某处开始全为 1：这时 x 中含有 a_2 作为其一段；

3) x 中含有无穷多个 0 及无穷多个 1，这时，我们取定某个 $x_k=1$，又分下列三种情形：

i）在 x_k 后有相邻两项均为 0（即有“00”），取 x_k 后的第一个“00”，则 x 中含有某 $100\beta_i$ 作为其一段；

ii）在 x_k 后无“00”，但有相邻两项均为 1，取这个“11”后的第一个 0，则 x 中含有某 1101 γ_j 为其一段；

iii）在 x_k 后既无“00”，亦无“11”，因而 x 中含有 a_3 作为其一段．

综上述，A 为基本集．

构造法也可用于解答计数问题：把计数的对象全部构造出来，则其个数不难求出．

例 3 如果一个多项式的所有系数均等于 0，1，2 或 3，则称为容许的．设 n 是自然数，求满足条件 $P(2)=n$ 的容许多项式 $P(x)$ 的个数．

解：如果我们先探求容许多项式的一般表达式，就能构造出满足 $P(2)=n$ 的全体容许多项式，而这样的多项式的个数也就可以求得了．

首先我们注意到：0，1，2，3 可以用惟一的方式写成 $a+2b$ 的形式，其中 a，b 的值为 0 或 1．于是 m 次容许多项式的一般表达式为

$$P(x)=\sum_{i=0}^{m}(a_i+2b_i)x^i . \quad ①$$

（a_i、$b_i=0$ 或 1，$i=0$，1，…，m）

这时，条件 $P(2)=n$ 成为：

$$\sum_{i=0}^{m}(a_i+2b_i)\times 2^i=\sum_{i=0}^{m}a_i\times 2^i+2\sum_{i=0}^{m}b_i\times 2^i=n . \quad ②$$

（a_i，$b_i=0$ 或 1，$i=0$，1，…，m）

若令 $A=\sum_{i=0}^{m}a_i\times 2^i, B=\sum_{i=0}^{m}b_i\times 2^i$，则②可写为

$$A+2B=n\ .\qquad ③$$

现在不难构造满足条件 $P(2)=n$ 的容许多项式：取自然数 A、B 满足③，将 A、B 写成2进制表示的形式：$A=\sum_{i=0}^{m}a_i\times 2^i, B=\sum_{i=0}^{m}b_i\times 2^i$（$a_i$、$b_i=0$ 或 1，$i=0, 1, \cdots, m$），则 $P(x)=\sum_{i=0}^{m}(a_i+2b_i)x^i$ 使 $P(2)=n$．因为 $A\geqslant 0$，由③得 $0\leqslant B\leqslant\left[\frac{n}{2}\right]$，故 A、B 的取法有 $\left[\frac{n}{2}\right]+1$ 种，因而使 $P(2)=n$ 的容许多项式有 $\left[\frac{n}{2}\right]+1$ 个．

几何中的作图问题也是一类构造性的问题，这类问题的一般提法是：给定若干已知元素（线段、角、定点等等），要求有限次地使用无刻度的直尺和圆规作出合乎某种条件的图形．

例4 任给 $\triangle A_1A_2A_3$，其三边之长为

$$a_1=A_2A_3,\ a_2=A_1A_3,\ a_3=A_1A_2\ .\qquad ①$$

F 为平面上的一个已知点．试从 F 出发在平面上引三条线段 FF_1、FF_2、FF_3，使其长度分别为

$$FF_i=a_i,\qquad ②$$

而线段的另外三个端点组成等边三角形 $F_1F_2F_3$．

解：我们给出两种构造的方法．

构造Ⅰ 设 $\triangle A_1A_2A_3$ 的内角为 A_1，A_2，A_3，由 $A_1+A_2+A_3=180°$，可得

$$(A_1-60°)+(A_2-60°)+(A_3-60°)=0\ .\qquad ③$$

不妨设 $A_1\geqslant A_2\geqslant A_3$．则

$$A_1-60°\geqslant 0,\ A_3-60°\leqslant 0\ .\qquad ④$$

在平面上沿顺时针方向作 $\angle F_2FF_3=A_1-60°$，使

$$FF_2 = a_2,\ FF_3 = a_3. \qquad ⑤$$

再以 FF_3 为边作$\angle F_3FF_1 = A_2 - 60°$（当 $A_2 - 60° \geqslant 0$ 时，$\angle F_3FF_1$ 沿顺时针方向，否则沿逆时针方向），使

$$FF_1 = a_1. \qquad ⑥$$

这时，由③可知$\angle F_1FF_2 = A_3 - 60°$且沿逆时针方向．在$\triangle F_2FF_3$中，以△表此三角形的面积，则由余弦定理得

$$\begin{aligned} F_2F_3^2 &= a_2^2 + a_3^2 - 2a_2a_3\cos(A_1 - 60°) \\ &= a_2^2 + a_3^2 - 2a_2a_3(\cos A_1\cos 60° + \sin A_1\sin 60°) \\ &= a_2^2 + a_3^2 - \frac{1}{2}\cdot 2a_2a_3\cos A_1 - 2\sqrt{3}\cdot\frac{1}{2}a_2a_3\sin A_1 \\ &= a_2^2 + a_3^2 - \frac{1}{2}(a_2^2 + a_3^2 - a_1^2) - 2\sqrt{3}\triangle \\ &= \frac{1}{2}(a_1^2 + a_2^2 + a_3^2 - 4\sqrt{3}\triangle). \end{aligned} \qquad ⑦$$

同理可得

$$F_1F_2^2 = F_1F_3^2 = \frac{1}{2}(a_1^2 + a_2^2 + a_3^2 - 4\sqrt{3}\triangle). \qquad ⑧$$

故$\triangle F_1F_2F_3$是边长为$\left[\frac{1}{2}(a_1^2 + a_2^2 + a_3^2 - 4\sqrt{3}\triangle)\right]^{\frac{1}{2}}$的等边三角形，构造完毕．

在构造Ⅰ中，当$\triangle A_1A_2A_3$为等边三角形时，$A_1 = A_2 = A_3 = 60°$，这时 F_1，F_2，F_3 重合为一点，因而$\triangle F_1F_2F_3$退化为“零三角形”．

构造Ⅱ　由 $A_1 + A_2 + A_3 = 180°$可知

$$(A_1 + 60°) + (A_2 + 60°) + (A_3 + 60°) = 360°. \qquad ⑨$$

在平面上沿逆时针方向作

$$\angle F_2FF_3 = A_1 + 60°,\ \angle F_3FF_1 = A_2 + 60°, \qquad ⑩$$

使 $FF_2 = a_2$，$FF_3 = a_1$，则由⑨可知

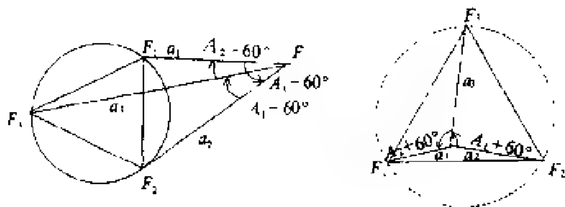


图 4 7

$$\angle F_1FF_2 = A_3+60°. \qquad ⑪$$

在$\triangle F_2FF_1$中，由余弦定理得

$$\begin{aligned}F_2F_3^2 &= a_2^2+a_3^2-2a_2a_3\cos(A_1+60°)\\ &= a_2^2+a_3^2-2a_2a_3(\cos A_1\cos 60°-\sin A_1\sin 60°)\\ &= a_2^2+a_3^2-\frac{1}{2}\cdot 2a_2a_3\cos A_1+2\sqrt{3}\cdot\frac{1}{2}a_2a_3\sin A_1\\ &= a_2^2+a_3^2-\frac{1}{2}(a_2^2+a_3^2-a_1^2)+2\sqrt{3}\triangle\\ &= \frac{1}{2}(a_1^2+a_2^2+a_3^2+4\sqrt{3}\triangle). \qquad ⑫\end{aligned}$$

同样可得

$$F_1F_2^2=F_1F_3^2=\frac{1}{2}(a_1^2+a_2^2+a_3^2+4\sqrt{3}\triangle). \qquad ⑬$$

于是 $\Delta F_1F_2F_3$ 是边长为 $\left[\frac{1}{2}(a_1^2+a_2^2+a_3^2+4\sqrt{3}\triangle)\right]^{\frac{1}{2}}$ 的等边三角形.

2. 构造法是证明存在性的直接方法

关于用构造方法解决数学中的存在性问题，上节中已作过讨论．在数学解题中亦有许多涉及存在性的问题，这些问题当然也可以用直接构造的方法来解答.

例1 求证：非质数亦非两质数之和的自然数有无穷多个.

证明：考察形如 $30m+5$ 的自然数（其中 m 为自然数），这样的数有无穷多个，每个均大于 5 且含有因数 5，故都不是质数．因为两个奇质数的和是偶数，而 $30m+5$ 是奇数，故不会是两个奇质数之和．惟一的偶质数是 2，但

$$(30m+5)-2=3(10m+1) \quad ①$$

不是质数，故 $30m+5$ 亦不会是 2 与一个奇质数之和．于是我们构造了无穷多个自然数，它们都不是质数且非两质数之和．

注　我们构造了无穷多个非两质数之和的合数，但我们没有（也不必）构造出全部这样的数．例如 57 不是 $30m+5$ 型的数，但 57 也是非两质数之和的合数．容易看出，上述问题等价于命题"差为 2 的合数对有无穷多对"，而此命题成立的一个充分条件是"对任意的自然数 n，存在相邻的 n 个自然数，它们都是合数"．后者亦可用构造法证明，因为

$$(n+1)!+2,(n+1)!+3,\cdots,(n+1)!+(n+1) \quad ②$$

恰是相邻的 n 个合数．

例 2　（格点与有理格点）

在平面上引入直角坐标系以后，坐标为整数的点称为格点．在数学的许多分支中，格点起着十分重要的作用．在 19 世纪时，几何数论的奠基人闵可夫斯基曾利用格点给出关于整数的许多优美的性质．

1957 年，史泰因豪斯提出了关于平面格点的一个问题，谢尔品斯基给出了这个问题的解答．所提的问题是："是否对任意自然数 n，都存在一个圆 C_n，C_n 中恰有 n 个格点．"

问题的答案是肯定的，谢尔品斯基用构造法证明了这一结论．

为了证明对任意自然数 n 存在恰含 n 个格点的圆 C_n，只要在平面上找出一点 P，P 到所有平面格点的距离均不相等．

谢尔品斯基构造了这样的一点，它就是 $P\left(\sqrt{2},\ \frac{1}{3}\right)$．我们证明 $P\left(\sqrt{2},\ \frac{1}{3}\right)$ 到所有平面格点的距离均不相等．

设 $A(a,b)$，$B(c,d)$ 是两个不同的格点，其中 a，b，c，d 为整数，且 $a\neq c$，$b\neq d$ 中至少有一个成立．若 $PA=PB$，即有

$$(\sqrt{2}-a)^2+\left(\frac{1}{3}-b\right)^2=(\sqrt{2}-c)^2+\left(\frac{1}{3}-d\right)^2, \qquad ①$$

则有

$$a^2+b^2-c^2-d^2-\frac{2}{3}b+\frac{2}{3}d=2\sqrt{2}(a-c). \qquad ②$$

但 a，b，c，d 均为整数，故②的左边为有理数，因而右边亦然，于是必有

$$2\sqrt{2}(a-c)=0, a=c. \qquad ③$$

代入②得

$$b^2-d^2-\frac{2}{3}(b-d)=0. \qquad ④$$

由于 $a=c$，故 $b\neq d$，由④得

$$b+d=\frac{2}{3}. \qquad ⑤$$

但⑤是不可能的，因为 b，d 是整数．于是我们证明了 P 到平面内任何两个格点的距离均不相等．

现设 n 为任意自然数，取正方形 Q_{ABCD}，其顶点为 $A(-(n+1),-(n+1))$，$B((n+1),-(n+1))$，$C((n+1),(n+1))$，$D(-(n+1),(n+1))$，则 Q 的边界及内部有 $(2n+1)^2>n$ 个格点且 $P\left(\sqrt{2},\ \frac{1}{3}\right)\in Q$．由于 P 到这些格点的距离均不相等，将其按大小排列为

$$0<r_1<r_2<\cdots<r_n<r_{n+1}<\cdots<r_{(2n+1)^2}. \qquad ⑥$$

取 $r=\frac{1}{2}(r_n+r_{n+1})$，则 $r_n<r<r_{n+1}$．记 C_n 为以 P 为圆心，r 为半径的圆，则 C_n 中恰含 n 个格点．

史泰因豪斯所提的存在性问题由谢尔品斯基的精巧的构造而得到肯定的答案．斯科恩伯格很欣赏谢尔品斯基的构造，他把坐标为有理数的点称为"有理格点"，进而提出："平面上到所有有理格点距离均不相等的点组成的点集是什么?" 斯科恩伯格研究了这个问题，并且给出了这个点集的一种刻画．

首先，如果点 $P(x_1,x_2)$ 到有理格点 $A(p_1,p_2)$、$B(q_1,q_2)$ 的距离相等，则有

$$(x_1-p_1)^2+(x_2-p_2)^2=(x_1-q_1)^2+(x_2-q_2)^2 . \quad ⑦$$

于是可得

$$2(p_1-q_1)x_1+2(p_2-q_2)x_2 - [(p_1^2+p_2^2)-(q_1^2+q_2^2)]=0 . \quad ⑧$$

因为 A，B 是不同的两点，故 $p_1-q_1\neq0$ 与 $p_2-q_2\neq0$ 中至少有一个成立，令

$$a_1=2(p_1-q_1),a_2=2(p_2-q_2),$$
$$a_3=-[(p_1^2+p_2^2)-(q_1^2+q_2^2)], \quad ⑨$$

则点 P 在直线

$$a_1x_1+a_2x_2+a_3=0 \text{ 上}. \quad ⑩$$

（a_1，a_2，a_3 为有理数，a_1，a_2 不全为 0）

由此得知：如果一点不在任何一条形如⑪的直线上，则这点到所有有理格点的距离均不相等．这是决定一点到所有有理格点的距离均不相等的一个充分条件．

其次，我们来考察上述条件的必要性．为证必要性，只需指出，对形如⑩的直线上的每一点，都存在一对有理格点，这点到这对有理格点的距离相等．斯科恩伯格用构造法证明了这

一对有理格点的存在性：因为 a_1，a_2 不全为 0，不妨设 $a_1 \neq 0$．则 $M\left(-\frac{a_3}{a_1}, 0\right)$在直线⑩上；取点 A，B 为有理格点：

$$A\left(-\frac{a_3}{a_1}+a_1, a_2\right), B\left(-\frac{a_3}{a_1}-a_1, -a_2\right),$$

则易知 M 是 AB 的中点；又过 A、B 的直线是

$$a_2x_1-a_1x_2+\frac{a_2a_3}{a_1}=0. \quad ⑪$$

故知直线⑩、⑪互相垂直，因而直线⑩是线段 AB 的垂直平分线，⑩上的每个点都与有理格点 A，B 等距．必要性由此得证．

3、递归构造是一种能行的构造方法

上面各例中的构造都是直接将对象构造出来，还有一类常见的构造方法是递归构造，这是一种能行的构造方法．递归构造法的理论依据是数学归纳法原理：如果要求构造的一类对象中含有自然数参数 n，则我们先对 n 较小时的情形作出构造(这是直接构造)，然后指出一种方法（也是直接构造的方法，故递归构造以直接构造为基础)，在假定对参数 n 所要求构造的对象已经完成的基础上，构造出对参数$(n+1)$所要求构造的对象．于是，根据数学归纳法原理，我们对任意的自然数 n 完成了所需要的构造．递归构造法是一种由简而繁，以简驭繁，逐次递进的构造方法，每一次构造都以前一次已完成的构造为基础，按照一种确定的模式进行．

递归构造法广泛地应用于各个领域、各种类型的问题中，请看下面的例子．

例 1 试将 1，2，…，n 排成一行，使其中任两数的算术平均值不等于排在这两数之间的任何数．

解：先设 n 是 2 的幂：$n=2^m$，当 $m=1$ 时，将 1、2 排

列成 12；当 $m=2$ 时，将 1、2、3、4 排列成 1324，均是合乎要求的排列．现设 1，2，…，2^m 已按要求排列成

$$a_1, a_2, \cdots, a_{2^m}, \quad ①$$

则易知

$$2a_1, 2a_2, \cdots, 2a_{2^m};$$
$$2a_1-1, 2a_2-1, \cdots, 2a_{2^m}-1 \quad ②$$

恰是 1，2，…，2^m，2^m+1，…，2^{m+1} 的一个合乎要求的排列．故对任何 2 的幂 2^m，均可构造 1，2，…，2^m 的合乎要求的排列．对于任意自然数 n，存在自然数 m，使 $2^{m-1}<n\leqslant 2^m$，我们先将 1，2，…，2^m 按要求排成一行，然后于其中划掉 $n+1$，$n+2$，…，2^m，剩下的即是 1，2，…，n 的合乎要求的排列．

例 2 证明：对于任意自然数 n，用数字 1 和 2 可以构造 2^{n+1} 个 2^n 位数，使其中的每两个数中至少有 2^{n-1} 个相应数位上的数字不相同．

证明：当 $n=1$ 时，四个二位数 11，22，12，21 中每两个都至少有一个相同数位上的数字不相同．现设我们已用 1 和 2 构造了 2^{n+1} 个 2^n 位数合乎题设的要求，用 A_n 记这 2^{n+1} 个数组成的集合．

对每个 $a\in A_n$，设 a' 表示将 a 中的 1 换成 2，2 换成 1 后所得的数，aa、aa' 分别表示 a 和 a、a 和 a' 连接而成的数，令

$$A_{n+1}=\{aa, aa' \mid a\in A_n\}, \quad ①$$

则得知 A_{n+1} 中共有 2^{n+2} 个由数字 1 和 2 组成的 2^{n+1} 位数，每两个数中至少有 2^n 个相应数位上的数字不相同．

例 3 求证：对任意自然数 k，关于 a，b，c 的方程 $a^2+b^2=c^k$ 均有正整数解．

证明：首先我们注意下述事实："若 M 为两个不同的自然数的平方和，N 为两个自然数的平方和，则 MN 亦为两个自然数的平方和．"事实上，设 $M=m^2+n^2$ $(m\neq n)$，$N=x^2+y^2$，则

$$\begin{aligned}MN&=(mx+ny)^2+(my-nx)^2\\&=(mx-ny)^2+(my+nx)^2,\end{aligned}\qquad ①$$

此处 $mx-ny$，$my-nx$ 不能同时为 0，否则 $m=n$，$x=y$，与所设矛盾．

其次，我们用递归的方法对任意的 k 构造方程的解．当 $k=1$ 时，命题是平凡的；当 $k=2$ 时，方程 $a^2+b^2=c^2$ 有正整数解(3，4，5)；现设方程 $a^2+b^2=c^{2r}$ $(r\geqslant1)$ 有正整数解，则由前述，$c^2\cdot c^{2r}=c^{2(r+1)}$ 亦为两自然数的平方和，即方程 $a^2+b^2=c^{2(r+1)}$ 有正整数解．故 k 为偶数时，用递归构造法恒可以得出方程的正整数解．当 k 为奇数时，由于已证方程 $a^2+b^2=c^{2k}$ 有正整数解 (p,q,r)，故方程 $a^2+b^2=c^k$ 有正整数解 (p,q,r^2)．故对任意自然数 k，$a^2+b^2=c^k$ 恒有正整数解．

例 4　求证：对任意自然数 $t\geqslant0$，多项式 $(x+1)^{2t+1}-x^{2t+1}$ 可以写成 $x(x+1)$ 的 t 次多项式，即有

$$(x+1)^{2t+1}-x^{2t+1}=f_t(x(x+1)),\qquad ①$$

其中 $f_t(u)$ 为 u 的 t 次多项式．

证明：当 $t=0$ 时，$(x+1)-x=1$，故 $f_0(u)=1$．

当 $t=1$ 时，$(x+1)^3-x^3=3x(x+1)+1$，故 $f_1(u)=3u+1$．

设对 $t\geqslant1$，已有①成立，则由

$$\begin{aligned}&(x+1)^{2t+3}-x^{2t+3}\\&=((x+1)^{2t+1}-x^{2t+1})((x+1)^2+x^2)+x^2(x+1)^2\,.\end{aligned}$$

$$((x+1)^{2t+1}-x^{2t+1}),$$

及

$$(x+1)^2+x^2=2x(x+1)+1,$$

可知

$$f_{t+1}(u)=(2u+1)f_t(u)+u^2f_{t-1}(u),$$

于是对任意自然数 $t\geqslant 0$, $f_t(u)$ 可由

$$\begin{cases} f_{t+1}(u)=(2u+1)f_t(u)+u^2f_{t-1}(u)(t\geqslant 1), \\ f_0(u)=1, f_1(u)=3u+1 \end{cases}$$

递归地构造.

4. 构造方法、策略、算法

有 10 箱肥皂, 其中 9 箱中的肥皂每块重 100 克, 但有一箱是次品, 每块重 99 克. 单看外表, 不能将次品辨认出来, 但是可用秤称. 请你设计一种方法, 用秤称一次即找出哪箱次品.

这个问题的困难在于只许用秤称一次, 这宝贵的一次必须慎用, 用秤之前必有所考虑. 我们可采用下面的方法: 先将 10 箱肥皂用 1, 2, …, 9, 10 编号, 然后从编号为 i 的箱中取出 i 块肥皂($i=1, 2, \cdots, 9, 10$), 共得 55 块肥皂, 若每箱都不是次品, 则不用秤也知道 55 块肥皂的总重量是 5500 克, 因为有次品, 故总重量不足 5500 克. 用秤一次称出 55 块肥皂的总重量为 x, 则 $(5500-x)$ 就是那箱次品肥皂的编号, 于是次品找出.

本题也是一种构造, 但不是构造具体的对象, 而是构造一种方法, 数学解题中我们也经常遇到需要构造方法、策略、算法的情况.

例 1 设有三堆棋子, 一次"操作"是指从某两堆棋子中各取出一枚棋子放入另一堆, 求证: 经过有限次操作可将 3 堆棋

子并为1堆的充要条件是，3堆中有2堆棋子的数目对模3同余．

证明：必要性．

设3堆中各堆所含棋子的数目分别为 x，y，z，从第1、第2堆中各取一枚棋子放入第3堆后，各堆棋子的数目变成 x'，y'，z'，其中

$$x'=x-1,\ y'=y-1,\ z'=z+2,$$

易知

$$x'-y'=(x-1)-(y-1)=x-y,$$
$$z'-x'=(z+2)-(x-1)=(z-x)+3,$$
$$y'-z'=(y-1)-(z+2)=(y-z)-3.$$

故进行一次操作后，各堆棋子的数目之差或保持不变，或增减3，但在操作过程中各堆棋子的数目对于模3的同余关系保持不变．

若经过有限次操作后3堆棋子并为1堆，即3堆中有两堆棋子的数目为0．它们对模3同余，故开始时的3堆棋子中必有两堆棋子的数目对模3同余．

充分性．

设3堆棋子中所含棋子的数目各为 x，y，z，且不妨设 $x-y=3k$．我们构造性地指出将3堆棋子经有限次操作并为1堆的具体方法，从而得证充分性．

首先，从两堆棋子中各取一枚并入另一堆，这是一次操作，将同样的操作重复若干次，则我们允许"从两堆棋子中各取相同数目的棋子并入另一堆"．

其次，如果我们先从第1、第2堆棋子中各取一枚并入第3堆，再从第1、第3堆棋子中各取一枚并入第2堆，则在此过程中棋子数目的变化为

$$(x, y, z) \longrightarrow (x-1, y-1, z+2) \longrightarrow$$
$$(x-2, y+1, z+1)$$

故两次操作的结果相当于"从1堆棋子中取出两枚分给另两堆各一枚"；将此过程重复若干次，则我们允许"从某堆棋子中取出偶数枚棋子，平分给另两堆"。

由于 $x-y=3k$，故 $x=y+3k$．我们先从第1堆棋子中取出 $2k$ 枚平分给第2、第3堆各 k 枚，则各堆棋子的数目变成

$$(x-2k, y+k, z+k)$$

但 $x-y=3k$，故 $x-2k=y+k=s$，即第1、第2堆中棋子的数目均等于 s；然后我们从第1、第2堆中各取 s 枚棋子并入第3堆，则第1、第2堆消失，只剩下第3堆，即所有棋子并为1堆．

数学中有一类博弈问题，这类问题的目标就是构造取胜的策略．例如，在边长为 $2n$ 的棋盘的左上角的一格中置一枚棋子，甲、乙两人对弈，轮流移动棋子，每次可将棋子移入相邻(即有一条公共边)的方格中，但不能移到棋子曾经到过的方格，谁如果面临无法走子的情况(棋子所在方格的相邻方格都是棋子曾经到过的方格)则告失败．若甲先走，则可以设计一种策略使甲必然取胜．事实上，因为棋盘的边长是偶数，故甲总可以默默地将棋盘划分为一些 1×2 的小长方形，甲走过一步后，棋子已占满一个长方形，故乙必将棋子移入一个新的(即棋子未曾进入的)长方形中，而甲可将棋子移入同一长方形的另一格．这样，乙每次都必将棋子移入新的长方形，而甲则将其移入此长方形的另一格．故甲永远有地方可走，因而不会失败，但棋盘有限，故乙必败而甲胜．

例2 桌上有三堆火柴，甲、乙对弈，两人轮流取火柴，

但每次只能从一堆中取走若干根火柴(也可以取走一堆)．最后剩下的火柴归谁所有，谁就获胜．问是否能够构造一种策略使甲必胜?

解：设三堆火柴中火柴的数目分别是 x、y、z，将此三数均用二进制表示．并将这三个二进制数从右向左各数位对齐排成三行(左边的数位不足时可用0填补)．这样排定后，每列(即三数的相同数位)都有三个数字"0"或"1"．我们称这三个数组成的数组是"正则"的，如果排定以后每列中都含有偶数个(0个或2个)"1"．

我们注意下面的事实：若将"从一堆火柴中取走若干根或整堆火柴"称为一次操作，则

1）对正则数组进行一次操作，必得非正则数组．这是因为进行一次操作后，必定有一行的一个数字发生变化，因而这个数字所在的列中"1"增加或减少1个，即数组的正则性被破坏；

2）对非正则数组进行一次适当的操作，必可将其化为正则数组．例如，对非正则数组

6	5	4	3	2	1
1	0	0	1	0	1
0	1	1	0	0	1
0	0	1	0	0	0

我们将各列从右至左编号，则此数组的第3、5、6列均只含1个"1"．若将第1数中第6、第3列的"1"去掉，而将第5列的"0"变成"1"，则数组将化为正则数组．这只要从第1数中减去

$$100100-10000=10100$$

即可实现．一般情形可仿此证明之．

假设甲先操作，如果开始时三堆火柴的数目组成非正约数组，则甲进行适当的操作将其化为正则数组，这时乙不论怎样操作，都会将数组变成非正则数组，轮到甲时，又将其化为正则数组，如此下去，乙永远不会得到正则数组，故他不会取走最后剩下的火柴(因为(0，0，0)是正则数组)．所以乙不胜而甲获胜．

综上所述，当开始时三堆火柴中火柴的数目组成非正则数组，则先走的人有必胜策略．

5．用构造思想解一类极值问题

在数学解题中常常会遇到这样一类极值问题，其中所涉及的量的取值是离散的(例如，取值为自然数)，并且量与量之间并无明确的函数关系，而要求的极值往往是变量的临界值(变量处在临界值的一侧时某种性质具备或某种要求可以实现，处在另一侧时则否．临界值即分界点)，这时，我们所熟悉的一些求极值的常规方法无法施展，所以可考虑用构造思想来解决问题．

用构造思想解这类极值问题有两种可行的方法，今以求极大值为例说明之．

其一，用构造的方法确定临界值的界，然后进行理论证明．

我们先构造一个具体的例子说明变量 $z=k$ 时某种条件不能满足，因而临界值(即能使条件满足的最大值)必小于或等于 k；若能证明当 $z\leqslant k-1$ 时条件确能满足，则所求的最大值为 $k-1$．

其二，假定条件成立或要求已经实现，用估计的方法确定 $z\leqslant k$，然后构造具体的例子说明 $z=k$ 确能保证条件成立，则

临界值(极大值)为 k .

求极小值的方法与此类似，详见下面的例 1、例 2 . 有时，估计所得的界不可以达到，故不是临界值. 故需通过放缩，适当调整，见例 3 .

例 1 在 100×25 的长方形表格中每一格填入一个非负实数，第 i 行第 j 列中填入的数为 x_{ij}($i=1$, 2, …, 100; $j=1$, 2, …, 25)(如表 1). 然后将表 1 每列中的数按由大到小的次序从上到下重新排列为

$$x'_{1j}\geqslant x'_{2j}\geqslant\cdots\geqslant x'_{100j}\ (j=1,2,\cdots,25)\text{(如表 2)}.$$

表 1

$x_{1,1}$	$x_{1,2}$	…	$x_{1,25}$
$x_{2,1}$	$x_{2,2}$	…	$x_{2,25}$
…	…	…	…
$x_{100,1}$	$x_{100,2}$	…	$x_{100,25}$

表 2

$x'_{1,1}$	$x'_{1,2}$	…	$x'_{1,25}$
$x'_{2,1}$	$x'_{2,2}$	…	$x'_{2,25}$
…	…	…	…
$x'_{100,1}$	$x'_{100,2}$	…	$x'_{100,25}$

求最小的自然数 k，使得只要表 1 中填入的数满足

$\sum\limits_{j=1}^{25}x_{ij}\leqslant 1\quad(i=1,2,\cdots,100)$，则当 $i\geqslant k$ 时，在表 2 中就能保证

$$\sum_{j=1}^{25}x'_{ij}\leqslant 1$$

成立。

解：k 的最小值为 97 .

(1) 取 $x_{ij}=\begin{cases}0, & 4(j-1)+1\leqslant i\leqslant 4j;\\ \frac{1}{24}, & \text{其余的 } i.\end{cases}$ $(j=1, 2, \cdots, 25)$

这时 $\sum_{j=1}^{25} x_{ij}=0+24\times\frac{1}{24}=1$. $(i=1, 2, \cdots, 100)$

满足题设条件．重排后有

$$x'_{ij}=\begin{cases}\frac{1}{24}, & (1\leqslant i\leqslant 96);\\ 0, & (97\leqslant i\leqslant 100).\end{cases} \quad (j=1, 2, \cdots, 25)$$

这时 $\sum_{j=1}^{25} x'_{ij}=25\times\frac{1}{24}>1$. $(1\leqslant i\leqslant 96)$

故 k 的最小值$\geqslant 97$.

(2) 首先证明：表 1 中必有一行(设为第 r 行)的所有数

$x_{r,1}$，$x_{r,2}$，$\cdots$，$x_{r,25}$

必在重排后所得表 2 的前 97 行中都出现．

事实上，若上述结论不成立，则表 1 的每一行中至少有一个数不在表 2 的前 97 行中出现，即表 2 的前 97 行中至多共有表 1 中 $100\times 24=2400$ 个数，这与表 2 的前 97 行共有 $25\times 97=2425$ 个数矛盾．

其次，由重排要求知表 2 中每列的数从上到下是由大到小排列的，故当 $i\geqslant 97$ 时，$x'_{ij}\leqslant x'_{97,j}\leqslant x_{rj}(j=1, 2, \cdots, 25)$．

故当 $i\geqslant 97$ 时

$$\sum_{j=1}^{25} x'_{ij}\leqslant\sum_{j=1}^{25} x_{rj}\leqslant 1 .$$

综合(1)、(2)知 k 的最小值为 97 .

例 2 有 1000 张票，用 000，001，$\cdots$，998，999 编号，

有 100 个票箱，用 00，01，…，98，99 编号．一张票允许放入某个票箱，如果从票的编号的三个数字中划去某个数字之后，剩下的两个数字恰是这个票箱的编号．问用此法将所有的票全部装入票箱，最少要动用多少个票箱？

解：首先，我们估计要动用的票箱的数目的下界．因为编号为"000"的票划去编号中的任一个数字后剩下的两个数只能是"00"，因而这张票只能放入编号为 00 的票箱，即编号为 00 的票箱必定不空．同理，编号为 11，22，…，99 的票箱也不空．设 r_i（$i=0,1,\cdots,9$）是编号以 i 开头的不空的票箱的数目，则 $r_i\geqslant 1$，取

$$r_k=\min\{r_i, 0\leqslant i\leqslant 9\}=x.$$

设以 k 开头的 x（$=r_k$）个不空的票箱的编号为

$$kt_1,\ kt_2,\ \cdots,\ kt_x.$$

则当 $\{p,q\}\subseteq\{0,1,\cdots,9\}\setminus\{t_1,\cdots,t_x\}$ 时，编号为 kpq 的票只能放入以 pq 编号的票箱内．由于 p、q 各取 $(10-x)$ 个不同的值，故这样的票有 $(10-x)^2$ 张，它们占用了 $(10-x)^2$ 个票箱．另外，由所设，编号以 t_1、t_2、…、t_x 开头的不空的票箱各不少于 x 个，故又得到 x^2 个不空的票箱，于是不空的票箱的数目不少于 $(10-x)^2+x^2$ 个，易知 $(10-x)^2+x^2$ 的极小值为 50，故不空（即已动用）的票箱不少于 50 个．

其次，我们具体构造一种方法，用 50 个票箱装下全部 1000 张票，从而证明下界 50 是可以达到的．为此，将 0，1，…，9 等 10 个数字等分为两组（例如 0～4 为一组，5～9 为另一组），编号中的两个数字属于同一组的票箱组成的集合为 A，则易知集合 A 包含 50 个票箱．每张票上有三个数字，而所有的数字只分为两组，故由抽屉原则，三个数字中必有两个数字属于同一组．保留这两个数字而删去另一个数字，则此票

可装入集合为 A 中的一个票箱，于是 50 个票箱可按规则装下 1000 张票.

综上述，按规则装下 1000 张票，最少要用 50 个票箱.

有时，估计所得的界并不是可以达到的，还需要通过放缩，适当调整，如下例所示.

例 3 坐标平面上的直线称为正规的，若此直线与坐标轴或坐标轴角的平分线平行或重合. 问连接平面上 6 个点的直线中，最多能存在几条正规直线?

解：设通过平面上某 6 个点有 k 条正规直线. 用 n_i $(1\leqslant i\leqslant 6)$ 表示通过第 i 点的正规直线的数目；用 m_j $(1\leqslant j\leqslant k)$ 表示第 j 条正规直线上的已给点的数目. 令 $S=n_1+n_2+\cdots+n_6$，则 S 是通过各已给点的正规直线的总数，并且当第 j 条正规直线上有 m_j 个已给点时，此直线在 S 中重复计算了 m_j 次，因而亦有 $S=m_1+m_2+\cdots+m_k$，所以

$$n_1+n_2+\cdots+n_6=m_1+m_2+\cdots+m_k .$$

但过一点的正规直线最多 4 条，故

$$n_1+n_2+\cdots+n_6\leqslant 4\times 6=24 .$$

而每条正规直线至少通过两个已给点，故

$$m_1+m_2+\cdots+m_k\geqslant 2k .$$

于是有 $2k\leqslant 24$，而得 $k\leqslant 12$.

但 $k\neq 12$，否则

$$m_1+m_2+\cdots+m_{12}=n_1+n_2+\cdots+n_6=24,$$

而 $\quad m_j\geqslant 2,\ n_j\leqslant 4$.

故可得

$$m_1=m_2=\cdots=m_{12}=2,\ n_1=n_2=\cdots=n_6=4 ,$$

故每条直线恰过两个已给点，而每个已给点恰有 4 条正规直线通过.

作一个矩形，其边平行于坐标轴，而且包含了6个已给点于其中，平行移动矩形的各边使他们恰通过6个已给点中的至少一点，得到一个新的矩形，此矩形的边为通过已给点的正规直线．故每边上恰有两个已给点，这些点不能都在各边的内部，否则由 $2\times4=8>6$，将至少有8个已给点，与给定6个点相矛盾．所以必有此矩形的一个顶点为已给点中的一个，设为A，过A恰有4条正规直线，其中有两条重合于矩形的一组邻边，一条穿过矩形的内部，另一条则在矩形的外部，其上亦应恰含有两个已给点，但这不可能，因为所有的已给点都包含在矩形之中．

我们已证得 $k\leqslant11$；图4－8中我们构造了6个点及通过这些点的11条正规直线，故连接平面上6个点的直线中，最多可存在11条正规直线．

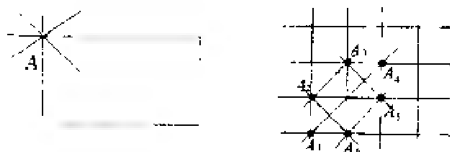


图4－8

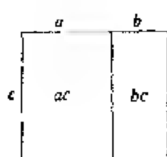
二、在解题过程中构造辅助工具

有些数学问题本身并不是构造性的，但在解题的过程中，根据问题的要求和性质恰当地构造辅助工具，往往有助于问题的解决．

1．构造图形

在小学算术中，图4－9用来说明乘法分配律；在初中代数中，图4－10用来说明两数和的平方公式．我们已经非常习

惯利用图形来帮助理解数学中的概念、定理、法则、公式，构造图形作为辅助工具，也是在数学解题中经常采用的一种方法。



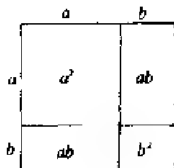


图 4－9　　　　图 4－10

构造图形可以作为数学解题的一种方法，是基于图形具有下面两个方面的特性：

第一，图形是一种几何结构，它可以将量的关系非常直观地显示出来，甚至超越了推理、论证，而将一切都凝聚于一图，使人一目了然。

例如，图 4－11 表示求自然数列前 n 项和的公式

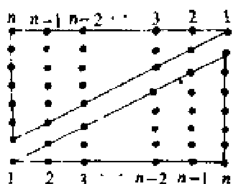


图 4－11

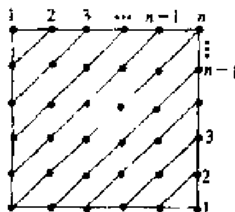


图 4－12

$$1+2+\cdots+n=\frac{1}{2}n(n+1),$$

推理已寓于图中。图 4－12 表示公式

$$1+2+\cdots+(n-1)+n+(n-1)+\cdots+2+1=n^2,$$

于是推出自然数列前 n 项和为：

$$1+2+\cdots+(n-1)+n=\frac{1}{2}(n^2+n)=\frac{1}{2}n(n+1).$$

若记 $S_1(n)=\sum_{k=1}^{n}k$，$S_3(n)=\sum_{k=1}^{n}k^3$，

则关系式 $S_3(n)=(S_1(n))^2$ 可用图 4－13 表示：

$$\begin{aligned}&(1+2+\cdots+n)^2\\&=1+2(1+2+1)+\\&\quad 3(1+2+3+2+1)+\cdots+\\&\quad n[1+2+\cdots+(n-1)+\\&\quad n+(n-1)+\cdots+2+1]\\&=1+2\times 2^2+3\times 3^2+\cdots+n\times n^2\\&=1^3+2^3+3^3+\cdots+n^3.\end{aligned}$$

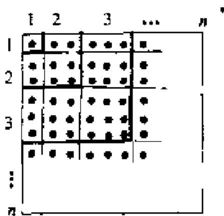


图 4－13

第二，图形本身具有量的性质：线有长度，面有面积，体有体积，因此，构造图形可以将一般的抽象的量，转化成比较具体的几何量，从几何量的关系中反映出抽象的量的关系，并且利用图形的划分、拼接，表示量的运算、相等、不等种种关系。

勾股定理是初等几何中最重要的定理，其公式 $a^2+b^2=c^2$ 可解释为：直角三角形立于斜边上的正方形的面积等于立于两直角边上的正方形的面积之和．因此，利用面积证明勾股定理就是一种十分自然的证法．当年欧几里得在《几何原本》中就是通过作图 4－14 中的辅助线利用面积关系证明勾股定理的，在他之后人们又设计了利用面积法证明勾股定理的种种方案，图 4－15 就是一个这样的方案，读者不妨一试．

证明两个图形面积相等，还有两种有趣的方法：剖分和拼补．如果两个多边形可以分别剖分为一些对应全等的多边形，则称它们“剖分相等”；如果两个多边形分别拼上一些对应全等

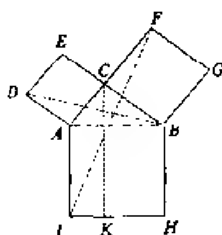


图 4－14

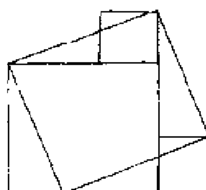


图 4－15

的多边形，变成两个全等的多边形，则称它们“拼补相等”．显然“剖分相等”成“拼补相等”的两个多边形面积相等．下

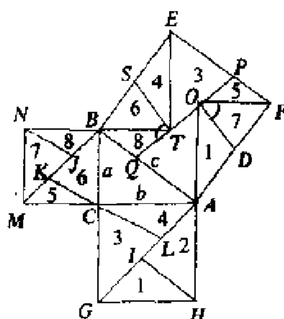


图 4－16 利用“剖分相等”证勾股定理

面的图 4－16 和图 4－17，就是利用剖分相等和拼补相等证明勾股定理的．

下面介绍用构造图形证明算术—几何平均值的方法：图 4 18利用面积的不等关系，其直观涵义几乎用不着说明；图4 19 利用线段的不等关系，该图所含的三角形是直角三角形，$\sqrt{ab}$乃根据定理“直角三角形斜边上的高是垂足分斜边所成的两线段的比例中项”求得．在这两个图形中还容易得到算术—几何平均值中等号成立的条件是 $a=b$．

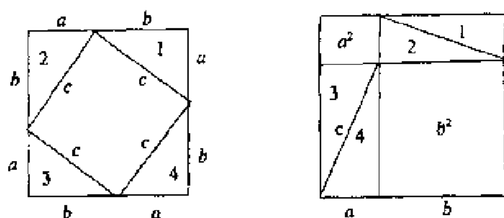


图 4－17　利用"拼补相等"证勾股定理

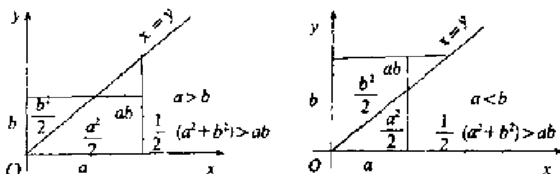


图 4－18

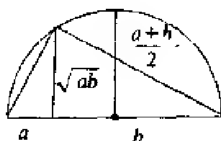


图 4－19

利用构造图形的方法解题，有时比用其他的方法更加直观、简捷、明了，而且不失严谨性．因此常常是一种美的享受，令人回味无穷！有一道流传其广的习题如下："已知图中的 $ABCD$，$DCEF$，$FEGH$ 都是正方形，求证 $\angle 1+\angle 2+\angle 3=90°$"（图 4－20）．这道题可用几何、三角、复数、向量等许多方法证明，但似乎都不及图 4－20 所示的"无字的证明"优美！

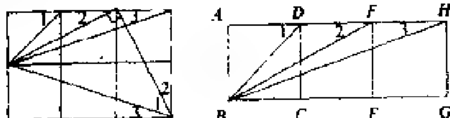


图 4－20

2. 构造函数

函数表示的是一般的量的关系，一个函数是许多具体的量的关系的概括，例如 $2+3$，$\frac{1}{2}+\frac{1}{3}$，$0.2+0.3$，$\sqrt{2}+\sqrt{3}$，…，可以概括为 $x+y$．在一些数学问题中，我们遇到一些具体的量，这些量具有相同的或类似的结构．如果我们构造一个函数来概括和反映这些量或量的关系，往往会有助于问题的解决．

例 1 已知

$$\begin{cases} \alpha^3-3\alpha^2+5\alpha=1, \\ \beta^3-3\beta^2+5\beta=5. \end{cases}$$

求 $\alpha+\beta$．

解：注意到两个已知等式的左边具有相同的结构，故可引入辅助函数

$$f(x)=x^3-3x^2+5x,$$

进而化成 $$f(x)=(x-1)^3+2(x-1)+3.$$

再引入函数

$$g(u)=u^3+2u,$$

则 $f(x)$、$g(x)$之间有关系

$$g(x-1)=f(x)-3.$$

易见 $g(u)$是单调上升的奇函数，而题中的条件变成

$$g(\alpha-1)=f(\alpha)-3=-2,$$

$$g(\beta-1)=f(\beta)-3=2.$$

由 $g(u)$的性质知 $\alpha-1$，$\beta-1$ 在 x 轴上关于原点对称，故有

$$(\alpha-1)+(\beta-1)=0,$$

由此得 $\alpha+\beta=2$．

例 2 设 $x_1,x_2,\cdots,x_n$ 为正数，求

$$\min_{\sum_{i=1}^{n} x_i=1} \max_{1\leqslant j\leqslant n}\left\{\frac{x_j}{1+x_1+\cdots+x_j}\right\},$$

这时 x_1，x_2，$\cdots$，x_n 的值是多少？

解：当

$$\sum_{i=1}^{n} x_i = 1,$$

$$\frac{x_1}{1+x_1}=\frac{x_2}{1+x_1+x_2}=\cdots=\frac{x_n}{1+x_1+\cdots+x_n}=d$$

时，易知

$$x_k=2^{\frac{k}{n}}\left(1-\frac{1}{2^{\frac{1}{n}}}\right),\ (k=1,\ 2,\ \cdots,\ n)$$

$$d=1-\frac{1}{2^{\frac{1}{n}}}.$$

我们证明

$$\min_{\sum_{i=1}^{n} x_i=1} \max_{1\leqslant j\leqslant n}\left\{\frac{x_j}{1+x_1+\cdots+x_j}\right\}=d.$$

若不然，则存在 x_1'，$\cdots$，x_n'，使

$$\sum_{i=1}^{n} x_i'=1,\ \max_{1\leqslant j\leqslant n}\left\{\frac{x_j'}{1+x_1'+\cdots x_j'}\right\}<d,$$

故有

$$\frac{x_1'}{1+x_1'}<\frac{x_1}{1+x_1},$$

$$\frac{x_2'}{1+x_1'+x_2'}<\frac{x_2}{1+x_1+x_2},$$

$$\cdots$$

$$\frac{x_n'}{1+x_1'+\cdots+x_n'}<\frac{x_n}{1+x_1+\cdots+x_n}.$$

引入辅助函数 $g(x)=\frac{x}{a+x}$ $(a>0,\ x\in[0,\ +\infty])$，则由

$$g(x)=1-\frac{a}{a+x},$$

可知$g(x)$在 $[0,\ +\infty]$ 上为增函数，于是由

$$\frac{x_1'}{1+x_1'}<\frac{x_1}{1+x_1}$$

得 $x_1'<x_1$，因而又有

$$\frac{x_2'}{(1+x_1')+x_2'}<\frac{x_2}{(1+x_1)+x_2}<\frac{x_2}{(1+x_1')+x_2},$$

故有 $x_2'<x_2$．一般地，若有 $x_1'<x_1$，$\cdots x_{j-1}'<x_{j-1}$，则由

$$\frac{x_j'}{1+x_1'+\cdots+x_{j-1}'+x_j'}<\frac{x_j}{(1+x_1+\cdots+x_{j-1})+x_j}$$

$$<\frac{x_j}{(1+x_1'+\cdots+x_{j-1}')+x_j},$$

可得 $x_j'<x_j$，于是有 $x_j'<x_j$ $(j=1、2\cdots,\ a)$，而

$$1=\sum_{i=1}^{n}x_i'<\sum_{i=1}^{n}x_i=1,$$

产生矛盾而得所欲证．

3．构造辅助元素

在平面几何中我们已经有了作辅助线的经验．通过作辅助线(平行线、垂线、连线、圆等等)，可以将问题中的已知条件联系起来，相互贯通，从而使各条件及条件与结论之间的关系显现出来．将作辅助线的思想推广到处理一般的数学问题，就产生了构造辅助元素的解题方法．

例 1 （关于排列的逆序总数）

设$(a_1,\ a_2,\ \cdots,\ a_n)$是$(1,\ 2,\ \cdots,\ n)$的一个排列，若 $i<j$ 而$a_i>a_j$（即一对数中较大的数排在较小的数的左边），则称 a_i，a_j 在此排列中构成一个逆序．一个排列中包含的逆

序的数目称为这个排列的逆序数．显然排列（1，2，…，n）的逆序数为0，而在排列（n，$n-1$，…，2，1）中，2排在1的左边，故（2，1）构成一个逆序，3排在2和1的左边，故（3，2），（3，1）构成两个逆序，…，n排在（$n-1$），…，2，1的左边，故$(n,n-1)$，$(n,n-2)$，…，$(n,2)$，$(n,1)$构成$(n-1)$个逆序，于是排列的逆序数为

$$1+2+\cdots+(n-1)=\frac{1}{2}n(n-1).$$

试求出全体n元排列(共$n!$个)的逆序数的总和(逆序总数)．

逐一求出每个n元排列的逆序数然后求和，显然不是求逆序总数的可行的方法．我们对每个排列（a_1，a_2，…，a_n），考虑用辅助元(a_n，a_{n-1}，…，a_2，a_1）去考察a_i，比a_i小的数共有a_i-1个，这些数或者在（a_1，a_2，…，a_n）中排在a_i的左边，或者在(a_n，a_{n-1}，…，a_2，a_1)中排在a_i的左边，故a_i在这两个排列中与比a_i小的数所构成的逆序共有a_i-1个，这样，我们就求出这两个排列的逆序数的和为

$$\sum_{i=1}^{n}(a_i-1)=\sum_{i=1}^{n}a_i-n=\frac{1}{2}n(n+1)-n$$

$$=\frac{1}{2}n(n-1).$$

由于$n!$个排列可配成$\frac{1}{2}(n!)$对，每对的逆序数之和都等于$\frac{1}{2}n(n-1)$，故逆序总数为$\frac{1}{4}n(n-1)(n!)$．

例2　在圆周上沿逆时针方向排列着n个数a_1，a_2，…，a_n．一次操作是指：将相邻的三个数x、y、z分别用$x+y$，$-y$，$y+z$代替．如果开始时圆周上排列的数依次是1，2，…，9，10，-1，-2，…，-9，-10，问能否通过有限次操

作，使圆周上的数的排列依次是 10，9，…，2，1，-10，-9，…，-2，-1．

解：设圆周上所排列的数依次是 a_1，a_2，…，a_n，且 $\sum_{i=1}^{n} a_i = 0$，又令 $a_{n+1}=a_1$，$a_{n+2}=a_2$，考察序列 a_1，a_2，…，a_n，a_{n+1}，a_{n+2}，对圆周上相邻三数的操作，就是对这个序列的相邻三项的相同的操作．

引进辅助序列 $\{b_i\}$，

$$b_i = \sum_{r=1}^{i} a_r = a_1 + a_2 + \cdots + a_i,$$

$$(i = 1, 2, \cdots, n+2)$$

$\{b_i\}$ 是序列$(a_1, a_2, \cdots, a_n, a_{n+1}, a_{n+2})$的部分和序列，由于 $\sum_{i=1}^{n} a_i = 0$，故 $b_{n+1} = b_1$，$b_{n+2} = b_2$．我们将 b_1，b_2，…，b_n 也沿逆时针方向依次排在一个圆周上．

设 $a_i = x$，$a_{i+1} = y$，$a_{i+2} = z$ $(1 \leqslant i \leqslant n)$是序列 $\{a_i\}$ 的相邻的三项，记 $d = \sum_{j=1}^{i-1} a_j$，则易知

$$b_i = d + x, \quad b_{i+1} = d + x + y,$$
$$b_{i+2} = d + x + y + z.$$

若对 a_i，a_{i+1}，a_{i+2}施行一次操作，即用 $x+y$，$-y$，$y+z$ 分别代替 x，y，z，则相应地 b_i，b_{i+1}，b_{i+2}变为

$$d + x + y, \quad d + x, \quad d + x + y + z.$$

因而我们看到：在圆周上对相邻三数 a_i，a_{i+1}，a_{i+2}进行一次操作，相当于在圆周上将相邻的两项 b_i，b_{i+1}交换次序；反之，交换 b_i，b_{i+1}的次序相当于对 a_i，a_{i+1}，a_{i+2}施行一次操作．

当原序列为$\{1, 2, \cdots, 9, 10, -1, -2, \cdots, -9, -10\}$时，写在圆周上的辅助序列（即原序列的部分和序列）是$\{1, 3, 6, 10, 15, 21, 28, 36, 45, 55, 54, 52, 49, 45, 40, 34, 27, 19, 10, 0\}$，要在圆周上通过对相邻三数的有限次操作将原序列化为$\{10, 9, \cdots, 2, 1, -10, -9, \cdots, -2, -1\}$，相当于要通过相邻两项的交换将写在圆周上的部分和序列化为新的部分和序列$\{10, 19, 27, 34, 40, 45, 49, 52, 54, 55, 45, 36, 28, 21, 15, 10, 6, 3, 1, 0\}$．由于两个序列所含的项完全相同，因而这是可以实现的．这说明通过问题中所规定的操作实现问题所要求的目的是可能的．

4．构造辅助命题

我们常常遇到这样的情形，为了证明一条定理，需要证明若干引理作为先导．在数学解题中也是这样：为了证明一个结论，我们常在更广阔的背景下考察这个结论，引进一些具有一般性的辅助命题，作为证明所需结论的依据．

例1 求证：在任意200个整数中必能找出100个数，使它们的和是100的倍数．

证明：用P_k表示命题："从$(2k-1)$个整数中恒可取出k个数，使其和为k的倍数．"

i）P_2，即"从三个数中恒可以取出两数，使其和是偶数"．事实上，任意三数中必有两数同奇偶性，而同奇偶的两数之和必为偶数．

ii）P_5，即"从9个数中恒可取出5个数，使其和是5的倍数"．

设9个数被5除的余数为

$$0 \leqslant r_1 \leqslant r_2 \leqslant \cdots \leqslant r_9 \leqslant 4 .$$

首先，若r_i（$i=1, \cdots, 9$）中有五个数相同，则结论成

立；若 r_i（$i=1$，…，9）中无任何三个数相同，则其中必同时含有 0，1，2，3，4，它们对应的五数之和是 5 的倍数，则结论亦成立。

其次，设 r_i（$i=1$，…，9）中恰有 3 个或恰有 4 个相同的数，设它们都等于 r。

1° 若 $r=0$，则在其余的 r_i（至少有五个数）中的五个，记之为 r_1'，r_2'，r_3'，r_4'，r_5'，考察和

r_1'，$r_1'+r_2'$，$r_1'+r_2'+r_3'$，$r_1'+r_2'+r_3'+r_4'$，$r_1'+r_2'+r_3'+r_4'+r_5'$。

如果其中有一个和数是 5 的倍数（它必不是 r_1'，因为 $r_1'\neq r=0$），故原数组中所对应的和是 5 的倍数；如果这些和都不是 5 的倍数，则其中必有两个和对模 5 同余（它们必不是相邻的两个和，否则其差即某个 $r_j'=0=r$）．这两个和数之差（也是若干 r_i' 之和）是 5 的倍数，故它在原数组中对应的和是 5 的倍数，这样在原数组中总有若干个数（至少两个、至多五个），它们不是 5 的倍数，但其和是 5 的倍数，又对应于 $r=0$ 的数（至少 3 个）是 5 的倍数，故原数组中必可取出五个数，其和是 5 的倍数。

2° 若 $r\neq0$，则于原数组中的 9 个数都各加上 $5-r$，则得到的新数组中恰有 3 个或恰有 4 个 $r_i=0$，由 1°，新数组中有 5 个数之和是 5 的倍数，将此 5 数还原，即各数均减去 $5-r$，因而和减少 $5(5-r)$，所剩即原数组中相应的五数之和是 5 的倍数。

综上所述，P_5 成立。

iii）今证明辅助命题："若 P_k，P_m 成立，则 P_{km} 亦成立"。

设有 $(2km-1)$ 个数，因 P_m 成立，故可于其中取出 m 个数，其和 S_1 是 m 的倍数；在剩下的数中取出 m 个数，其和

S_2 是 m 的倍数，这个手续可进行$(2k-1)$次，因为

$$(2km-1)-(2k-1)m=m-1>0.$$

由于$\frac{S_1}{m}$，$\frac{S_2}{m}$，…$\frac{S_{2k-1}}{m}$是 $2k-1$ 个整数，由 P_k 成立，可在其中取 k 个数，其和是 k 的倍数，即原数组中存 km 个数，其和是 km 的倍数。

iv）因为 $100=2\times2\times5\times5$，由 i）、ii），$P_2$，$P_5$ 成立，反复应用 iii），可知 $P_{2\times2}$，$P_{2\times2\times5}$，$P_{2\times2\times5\times5}$均成立，即 P_{100} 成立，因而在 $2\times100-1=199$ 个数中可取出 100 个数，其和是 100 的倍数，问题由此得证。

5．构造模型

构造模型作为辅助工具，也是一种常用的解题方法，因为模型具有下面的一些性质：

（1）直观性．一个数学问题总是涉及许多抽象的数学概念和概念之间的关系．为了理解概念和掌握它们之间的关系，我们总是用我们比较熟悉、比较具体的事物来构造适合问题的模型，概念及其相互关系在模型中反映出来，问题便有了真实、生动的形象，我们的想象和思考有了比较具体的对象，思维有了依托，这就是模型的直观性。

例如将 m 个不可辨的球放入 n 个可辨的盒子中，我们用“○”代表球，用$(n+1)$条短竖线“|”两两之间的间隔代表盒子，则一种放球的方法可视为 m 个“○”和$(n+1)$条“|”的一个排列(两端必排“|”)．于是容易得出放球的方法数为 C_{m+n-1}^{m}．

又如，斐波那契数列由递归关系

$$\begin{cases} f_{n+2}=f_{n+1}+f_n, & (n\geqslant0) \\ f_0=f_1=1 \end{cases}$$

定义，若令 $A=\begin{bmatrix}1&1\\1&0\end{bmatrix}$，则

$$A^0=E=\begin{bmatrix}1&0\\0&1\end{bmatrix}=\begin{bmatrix}f_1&f_0\\f_0&f_{-1}\end{bmatrix},\text{（定义 }f_{-1}=1\text{）}$$

$$A^1=A=\begin{bmatrix}1&1\\1&0\end{bmatrix}=\begin{bmatrix}f_2&f_1\\f_1&f_0\end{bmatrix}.$$

若有

$$A^n=\begin{bmatrix}f_{n+1}&f_n\\f_n&f_{n-1}\end{bmatrix},$$

则由递归关系式有

$$A^{n+1}=AA^n=\begin{bmatrix}1&1\\1&0\end{bmatrix}\times\begin{bmatrix}f_{n+1}&f_n\\f_n&f_{n-1}\end{bmatrix}$$

$$=\begin{bmatrix}f_{n+2}&f_{n+1}\\f_{n+1}&f_n\end{bmatrix}.$$

故对任意 n，均有 $A^n=\begin{bmatrix}f_{n+1}&f_n\\f_n&f_{n-1}\end{bmatrix}$，即矩阵序列 $\{A^n\}$ 中各项右上角元素组成斐波那契数列，这个模型称为斐波那契数列的矩阵表示．这种表示使用矩阵工具研究斐波那契数列的性质成为可能．

（2）可操作性．由于模型的直观性，构造模型的"材料"是我们比较熟悉的具体的事物，这样就便于对模型进行操作和变换．对模型的这些操作和变换，就是间接地对模型所反映的数学问题的操作与变换．

例如组合数 C_{m+n}^m 有下面的几何模型：如图 4－21，将坐标平面分成 1×1 的小方格，小方格的顶点的坐标为整数，称为平面格点．一质点从原点(0，0）出发，沿小方格的边向

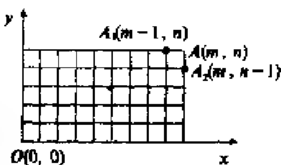


图 4－21

右或向上运动，最后到达格点(m, n)，C_{m+n}^{m}即表示质点按这种方式运动，从$(0, 0)$到(m, n)的所有可能的路径的数目．这个模型对于涉及组合数的许多问题非常有用，许多有趣的组合恒等式可以通过对模型的分析、变换而轻易地得到：

1．(杨辉恒等式)

$$C_{m+n}^{m}=C_{m+n-1}^{m}+C_{m+n-1}^{m-1}.$$

根据上面的模型，等式的左端是从$(0, 0)$到(m, n)的路径数；等号右端的第一项表示从$(0, 0)$到$(m-1, n)$的路径数，第二项为从$(0, 0)$到$(m, n-1)$的路径数，因$(m-1, n)$、$(m, n-1)$是从$(0, 0)$到(m, n)的必经的点，故从$(0, 0)$到(m, n)的路径数等于分别到$(m-1, n)$和$(m, n-1)$的路径之和，因而等式成立．

2．$C_{n+r+1}^{r}=C_{n+r}^{r}+C_{n+r-1}^{r-1}+C_{n+r-2}^{r-2}+\cdots+C_{n+1}^{1}+C_{n}^{0}$．

此式易从杨辉恒等式而得出，但根据模型观察则更加直观：等式左端表示从$(0, 0)$到$(n+1, r)$的路径的数目，如图4－22，用直线$x=n$截这些路径，直线与路径垂直相交的交点可为$(n, 0)$，$(n, 1)$，…，(n, r)，

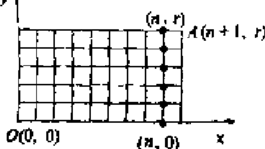


图4－22

故可根据交点的不同而将路径分为$(r+1)$类，与直线在点(n, i)垂直相交的路径的数目为$C_{n+i}^{i}(i=0, 1, \cdots, r)$，其和即右边各项之和，亦表示从$(0, 0)$到$(n+1, r)$的路径的总数．

前面给出了斐波那契数列的矩阵表示，所以矩阵可以成为研究斐波那契数列的工具，即斐波那契数列的许多性质可由矩阵的运算而得到．

例如，利用矩阵表示很容易证明下列等式

i) $f_0+f_1+\cdots+f_n=f_{n+2}-1$;

ii) $f_n=f_{k+1}f_{n-k}+f_kf_{n-k-1}$, $(n\geqslant k)$

特别地 $f_{2n+1}=f_{n+1}^2+f_n^2$.

iii) $f_{n-1}f_{n+1}-f_n^2=(-1)^n$. 由此可知相邻的斐波那契数互质.

事实上，当 $A=\begin{bmatrix}1&1\\1&0\end{bmatrix}$ 时，$(I-A)^{-1}=-A$. 故

$$I+A+\cdots+A^n=(I-A)^{-1}(I-A)(I+A+\cdots+A^n)$$
$$=(-A)(I-A^{n+1})=A^{n+2}-A.$$

比较上式两边矩阵右上角的元素，即得i).

由 $A^n=A^k\cdot A^{n-k}$，运用矩阵乘法将右方之积求出，然后比较两边矩阵的右上角的元素，即得ii).

由于 A 的行列式 $|A|=\begin{vmatrix}1&1\\1&0\end{vmatrix}=-1$,

故 $|A^n|=(|A|)^n=(-1)^n$.

即 $\begin{vmatrix}f_{n+1}&f_n\\f_n&f_{n-1}\end{vmatrix}=f_{n-1}f_{n+1}-f_n^2=(-1)^n$.

(3) 概括性. 一个好的模型，不仅适用于某个具体的问题，而且适用于形式相近、性质相同的一类或几类不同的问题. 也就是说，好的模型概括地反映了一类数学问题的共同的本质，因而可以有广泛的应用.

矩阵可以作为模型表示斐波那契数列，它也可以用来表示一般的线性递归数列. 事实上，在线性递归数列的研究中，矩阵早已成为一种有力而且有效的工具.

前面用“○”和“|”的排列解决放球问题的模型，也可以用来讨论方程

$$x_1+x_2+\cdots+x_n=m$$

的非负整数解的数目．因为求这个方程的非负整数解，相当于将排成一行的 m 个"○"用 $(n-1)$ 条"|"分开，左起第一条"|"的左边、相邻的两条"|"之间、最右的一条"|"的右边共 n 个间隔，每个间隔中所含"○"的个数依次作 x_1，x_2，…，x_n 的值，即得到方程的一组非负整数解．于是方程的非负整数解与 m 个"○"、$(n-1)$ 条"|"的排列相对应，解的数目等于排列数 C_{m+n-1}^{m}．如果我们要求这个方程的正整数解的数目，则仍然可以用这个模型，但每两条"|"中间必有"○"，亦即每两个"○"中至多只许插入一条"|"，故正整数解的组数即排列数为 C_{m-1}^{n-1}．

下面再举一个构造模型求解的典型例子．

例 1 （选举问题）

在一次选举中有甲、乙两名候选人，选举结果，甲得 m 票，乙得 n 票，$m \geq n$．若在唱票统计票数的过程中甲的票数总是领先于乙，问这样的唱票过程有多少种可能情况？

解：我们用一动点的运动来记录唱票过程．开始时，设动点处于坐标原点．若在唱票过程的某一时刻动点处于点 (x, y)，这时，若甲增加一票，则动点移到点 $(x+1, y+1)$；若乙增加一票，则移至 $(x+1, y-1)$．容易知道，某一时刻动点所在位置（点）的纵坐标，表示这一时刻甲与乙所获得的票数的差，横坐标表示开出的票的总数．而整个的唱票过程，即用连接原点和点 $(m+n, m-n)$ 的一条平面折线表示．在唱票的过程中甲恒领先于乙，即指动点在运动的中途不与 x 轴相遇．

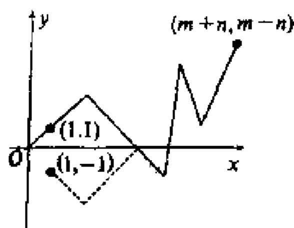


图 4－23

有了模型，即便于操作．我们知道，从点(0，0)到点$(m+n，m-n)$的路径的总数为C_{m+n-1}^{m-1}．但我们要求甲在唱票过程中领先，故第一票应为甲得，所以路径必过点(1，1)，即我们可以只考虑从点(1，1)到点$(m+n，m-n)$的路径．这样的路径共有C_{m+n-1}^{m-1}条，它们可分为两类：第一类不与x轴相遇，其余为另一类．我们要求的恰是第一类路径的数目，它等于路径的总数C_{m+n-1}^{m-1}减去从(1，1)出发到点$(m+n，m-n)$中途到达或穿越x轴的路径的数目．

为了求出与x轴相交的路径的数目，我们对这类路径作一个变换：取一条这样的路径，设C是这条路径与x轴的第一个交点，将路径上的(1，1)到C的一段(它在x轴的上方)关于x轴作轴对称变换，其图象是在x轴下方从(1，－1)到C的一条折线，这条折线加上原路径C后的一段，合成连接(1，－1)与$(m+n，m-n)$的一条折线；反之，每一条连接(1，－1)与$(m+n，m-n)$的折线，可还原成从(1，1)出发到$(m+n，m-n)$中途与x轴相遇的路径，两者之间的对应是一一对应，故从(1，1)出发到$(m+n，m-n)$而中途与x轴相遇的路径的数目，等于连接(1，－1)与$(m+n，m-n)$两点的折线的数目，后者等于C_{m+n-1}^{m}，于是求得从(0，0)到$(m+n，m-n)$而中途不与x轴相遇的路径的数目为

$$C_{m+n-1}^{m-1}-C_{m+n-1}^{m}=\frac{m-n}{m+n}C_{m+n}^{n}.$$这也就是原问题的答案．

我们指出，这个模型也适用于下面的“购票问题”：某剧院卖票的窗口前有$m+n$个人排队购票，其中m个人持有10元的钞票，n个人持有5元的钞票，票价每张5元，在开始售票时，售票员手中没有钱，要使售票的过程顺利进行而不因找零发生困难，问排队的方式有多少种可能情况．

只要注意排在前面任一段的顾客中持5元钞票人恒比持10元钞票的多，找零就不会发生困难，则购票问题与选举问题在本质上属相同的问题。

6. 构造例和反例

在数学中，为了深刻地认识概念或推进理论，常常需要构造各种各样的例子：构造一个例子以支持一个结论或构造一个反例以否定一个结论。优美的例子不仅作为理论的补充丰富了数学内容，而且有许多著名的例子在数学发展的过程中还起着重要的作用。

人们为着不同的目的而构造形形色色的例子。

为了证明合乎某种条件的对象的存在性，只要构造一个这种对象；为了证明合乎某种条件的对象存在无穷多个，只要列举无穷多个这种对象。为了证明存在平面无穷点集 S，使其中任意3点是一个面积为有理数的三角形的顶点，并且任意两点的距离为无理数，我们可构造下面的例子。

取平面抛物线 $y=x^2$ 上无穷多个点组成点集 $S=\{P_k\}$，$P_k(k,k^2)$，$k=1,2,3,\cdots$，于是 S 中任意两点的距离为

$$P_iP_j=\sqrt{(i-j)^2+(i^2-j^2)^2}=|i-j|\sqrt{1+(i+j)^2},$$

任意3点为顶点的三角形面积为

$$S_{\triangle P_iP_jP_k}=\frac{1}{2}\begin{vmatrix}1 & i & i^2\\ 1 & j & j^2\\ 1 & k & k^2\end{vmatrix}\text{的绝对值。}$$

显然 P_iP_j 为无理数且 $S_{\triangle P_iP_jP_k}$ 为有理数。

为了否定一个结论，只需举出一个反例，要否定命题“底数和指数都是无理数的幂是无理数”，由 $(\sqrt{2})^{\log_2 9}=3$ 立即可得。欧拉曾经想构造一个多项式 $f(x)$，使对任意的自然数 x，

$f(x)$都是质数．他构造了$f(x)=x^2+x+7491$，并且验证了$x=1, 2, \cdots, 10000$ 时，$f(x)$都是质数，但后来发现$f(72490)$不是质数，一个反例否定了欧拉的想法．

为了廓清概念、消除成见，也常借助于构造例子，狄里赫勒函数

$$D(x)=\begin{cases}1, & x\text{ 为有理数；}\\ 0, & x\text{ 为无理数，}\end{cases}$$

说明函数概念的广泛性，这个概念的本质在于数值对应．它还告诉我们，不要认为函数的图象一定是一条或一段曲线，也并非所有周期函数都有最小正周期（$D(x)$以任何非 0 有理数为周期，故无最小正周期）．

为了理解定理，分清条件的必要性和充分性，也需要构造正、反两个方面的例子．例如，图 4－24 说明两边和其中一边所对的角对应相等时，并不能保证两个三角形全等．又如图 4－25 中$\triangle ABC$为等边三角形，$\angle 1=\angle 2$，$DE=AC$，故知$\triangle ADC \cong \triangle DAE$，因而有$\angle E=\angle C$，于是在四边形 $ABDE$ 中，对边 $AB=AC=DE$，对角$\angle B=\angle C=\angle E$，但 $ABDE$ 不是平行四边形．这说明一组对边相等、一组对角相等的四边形，不一定是平行四边形．

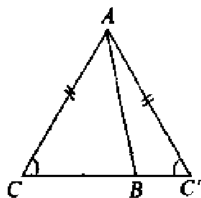


图 4－24

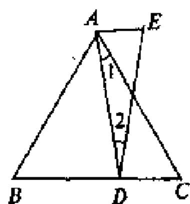


图 4－25

构造例子也常用于估计数量界限．若我们能够构造出一类

总体的 a 个个体，则关于这类总体的计数，其上界必不小于 a，其下界必不大于 a．若对某个量 X 得到估计式 $X \geqslant a$，又能够构造特殊情况作 $X=a$，则说明下界 a 可以达到，而所得的估计不可以改进．

著名数学家所构造的许多精彩的例子往往发人深省，它反映了概念的特殊属性和深刻的本质，把人们的认识引向直观无法达到的深度．黎曼函数

$$R(x)=\begin{cases}0, & x\ \text{为无理数}; \\ \dfrac{1}{q}, & x=\dfrac{p}{q},\ p、q\ \text{为互质自然数}.\end{cases}$$

当 x 为无理数时，$R(x)$ 连续；当 x 为有理数时，$R(x)$ 不连续，说明函数在连续点的任意近傍有间断点，而在间断点的任意近傍有连续点．范德瓦尔登、魏尔斯特拉斯分别构造了连续但处处不可导的函数的例子，消除了人们以为"连续函数除个别点外总是处处可导"的误解．皮亚诺曲线可以充塞一个正方形，它告诫人们原来曲线的定义失之过宽，因而有悖于常理，这些例子在数学史上是光辉夺目的！另外，如发现"正方形的边长与对角线不可通约"，是无理数产生的几何方面的原因．罗素悖论(其实也是一个例子)的提出促使数学基础研究的深入和三大学派的形成，它们对于推动数学的发展产生了很大的影响．

例 1 （n 阶点集问题）

平面上的一个点集，每两点之间的距离都相等，则此点集最多含有 3 个点，它们是正三角形的三个顶点；空间中的一个点集，每两点之间的距离都相等，则此点集最多含有 4 个点，它们是正四面体的 4 个顶点．

一般地，我们可以定义 n - 阶点集的概念：如果一个点集

中每两点之间的距离可以取 n 种不同的数值(这些数值不预先限定)，则称之为 n - 阶点集；含有的点数最大的 n - 阶点集，称为极大 n - 阶点集；若将点集中的点限制在同一直线上(同一平面上或三维空间中)，则称为直线上(平面上或空间中)的 n - 阶点集．依此定义，平面上的极大 1 - 阶点集含 3 个点，空间中的极大 1 - 阶点集含 4 个点．

直线上的 n - 阶点集很简单：直线上的极大 n - 阶点集含 $(n+1)$ 个点，它们是一条线段的两个端点及 $(n-1)$ 个 n 等分点(图 4 - 26)．若以 l_n 记直线上极大 n - 阶点集中点的数目，则有 $l_n = n+1$．

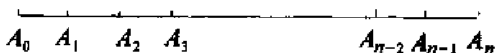


图 4 - 26

为了弄清楚平面上 n - 阶点集的结构及最大 n - 阶点集所含点数，我们先作一番探求．显然，正方形的 4 个顶点组成一个平面 2 - 阶点集，但不是极大 2 - 阶点集，因为正五边形的 5 个顶点也组成 2 - 阶点集．可以证明：平面上的极大 2 - 阶点集恰含有 5 个点．若以 P_n 记平面上极大 n - 阶点集中点的数目，则有

$$P_1 = 3;\ P_2 = 5\ .$$

于是我们猜想：

$$P_n = 2n+1\ .$$

每当我们提出猜想，我们总希望证实或否定．在作出一般证明之前，通常我们可构造实例以支持猜想，也可以构造一个反例来否定猜想．

考察正 $(2n+1)$ 边形的 $(2n+1)$ 个顶点，因为 $(2n+1)$ 边

形的边和对角线恰有 n 种不同长度，故由$(2n+1)$个顶点组成 n－阶点集．于是我们得到 P_n 的一个估计式：

$$P_n \geqslant 2n+1.$$

似乎可以猜想等号必成立．即正$(2n+1)$边形的$(2n+1)$个顶点组成极大 n－阶点集．

我们希望证明我们的猜想，然而未能成功．后来终于得到一个反例，否定了上述猜想．反例如下：[①] 将 144 个点置于 11×11 的正方形格点上，这些点之间有 71 种不同的距离：从图 4－27 之左上到右下考察每个 $i\times i$（$i=1, 2, \cdots, 11$）的正方形中各点之间的不同距离，易得：

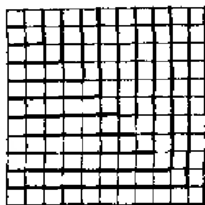


图 4－27

1．$\sqrt{2}$；

2．$\sqrt{5}$、$2\sqrt{2}$；

3．$\sqrt{10}$，$\sqrt{13}$，$3\sqrt{2}$；

4．$\sqrt{17}$，$2\sqrt{5}$，5，$4\sqrt{2}$；

5．$\sqrt{26}$，$\sqrt{29}$，$\sqrt{34}$，$\sqrt{41}$，$5\sqrt{2}$；

6．$\sqrt{37}$，$2\sqrt{10}$，$3\sqrt{5}$，$2\sqrt{13}$，$\sqrt{61}$，$6\sqrt{2}$；

7．$\sqrt{53}$，$\sqrt{58}$，$\sqrt{65}$，$\sqrt{74}$，$\sqrt{85}$，$7\sqrt{2}$；

8．$2\sqrt{17}$，$\sqrt{73}$，$4\sqrt{5}$，$\sqrt{89}$，10，$\sqrt{113}$，$8\sqrt{2}$；

9．$\sqrt{82}$，$3\sqrt{10}$，$\sqrt{97}$，$\sqrt{106}$，$\sqrt{117}$，$\sqrt{130}$，$\sqrt{145}$，$9\sqrt{2}$；

10．$\sqrt{101}$，$2\sqrt{26}$，$\sqrt{109}$，$2\sqrt{29}$，$5\sqrt{5}$，$2\sqrt{34}$，$\sqrt{149}$，$2\sqrt{41}$，$\sqrt{181}$，$10\sqrt{2}$；

11．$\sqrt{122}$，$\sqrt{130}$，$\sqrt{137}$，$\sqrt{146}$，$\sqrt{157}$，$\sqrt{170}$，

---

① 此例是陈计告诉作者的，谨致谢忱！

$\sqrt{185}$，$\sqrt{202}$，$\sqrt{221}$，$11\sqrt{2}$.
从而可知 $P_{71}\geqslant 144$，比猜想的值 $2\times 71+1=143$ 大．因此平面上的 n-阶（$n\geqslant 3$）点集的结构及 P_n 的值，尚是一个值得研究的大问题．

§4.3　算法化思想与机器证明

一、算法与算法化

在第二章中我们曾经指出：算法化思想是我国古代数学中一种最重要的数学思想，它与古希腊的公理化思想一起成为数学发展的两个重要源泉．在现代，由于电子计算机的出现和高度发展，科学计算、科学实验和逻辑推理已经成为发展数学和整个科学的三种最重要的手段，算法化的思想在推动数学乃至整个科学的发展方面已经并且必将产生越来越大的作用．

算法化思想是数学中一种重要的构造思想，这种思想的核心是构造算法．这就是说：在我们不能直接得出某个问题的解答的时候，可以设计或构造一种可行的步骤（计算公式或程序），使得能在有限步内得出问题的解，或者逐步逼近问题的解，或者能根据所得结果对所论对象的某些特性（如存在性、存在的个数等）作出正确的判断．

利用算法解题常有以下几种类型．

1．构造算法通过计算解决问题

（1）构造一个算法，经有限步计算得出问题的解．

例　我国古代数学著作《孙子算经》中有“物不知其数”一问，原文如下：“今有物不知其数，三三数之剩二，五五数之剩三，七七数之剩二，问物几何？”

在程大位所著《算法统宗》(1583)中，给出了如下的口诀："三人同行七十稀，五树梅花廿一枝，七子团圆正半月，除百零五便得知"。这是求解"物不知其数"问题的一个算法，其含义是：用 70 乘 3 除所得的余数，21 乘 7 除所得的余数，15 乘 7 除所得的余数，然后将所得三个乘积相加，若此和大于 105，则减去 105，若仍大于 105，则再减，最后得到一个小于 105 的自然数，它就是问题的解答。用此算法解"物不知其数"问题，用算式表示为：

$$70\times2+21\times3+15\times2=233,$$

$$233-105\times2=23.$$

故物之数为 23。

其他，如求两数的最大公约数的欧几里得算法，求线性方程组的解的高斯消元算法，求一元二次方程的根的公式算法等等，都属于这一类型。

(2)构造一个算法，经过计算，直接证明具有某种性质的对象是存在的。

例 证明：在 1，2，…，$\frac{3^n+1}{2}$中可以取出 2^n 个正整数，使其中的任意三个数不组成等差数列。

分析：设取出的 2^n 个正整数组成的集合为 A_n。为了探求算法，我们先从简单情形入手：

当 $n=1$ 时，$\frac{(3^1+1)}{2}=2$，故给定的数为 1，2，又 $2^1=2$，取 $A_1=\{1, 2\}$即合乎要求；

当 $n=2$ 时，$\frac{3^2+1}{2}=5$，故给定的数为 1，2，3，4，5，又 $2^2=4$，取 $A_2=\{1, 2, 4, 5\}$即合乎要求。

当 $n=3$ 时，我们将 A_2 中的 4 个数及其分别加上 3^2 后所

得的 4 个数共 8 数组成 A_3：

$$A_3=\{1,2,4,5;10,11,13,14\},$$

不难验证 A_3 合乎要求.

于是我们可以得到构造 A_n 的一个算法：

1）取 $A_1=\{1,2\}$，

2）若 $A_n=(n\geqslant 1)$已取定，则取

$$A_{n+1}=A_n\cup\{3^n+a \mid a\in A_n\}.$$

我们证明按此算法求得的 A_n 合乎要求.

首先，设 A_n 中有 a_n 个数，则可得算法

$$a_1=2;\ a_{n+1}=2a_n.\quad (n=1,2,\cdots)$$

于是易得 $a_n=2^n$，即 A_n 由 2^n 个正整数组成.

其次，用 b_n 表示 A_n 中的最大元，则可得求 b_n 的算法：

$$b_1=2;\ b_{n+1}=b_n+3^n.\quad (n\geqslant 1)$$

由此可得 $b_n=\frac{1}{2}(3^n+1)$，可见 A_n 是 $\left\{1,2,\cdots,\frac{3^n+1}{2}\right\}$的一个子集.

最后，注意到"三数 x、y、z 成等差数列，当且仅当对任意的 a，$a+x$，$a+y$，$a+z$ 成等差数列"，则用归纳法不难证明 A_n 中的任何三数均不成等差数列.

于是得到命题的证明.

读者或许已经发现，前面关于递归构造的例子，即属此种类型.

(3)构造一个算法，经过计算，间接求出原问题的解.

例 已知函数 f：$[0,+\infty)\to[0,+\infty)$满足

$$f(f(x))+f(x)=6x.$$

试求满足上述条件的一切 f.

分析：记 $f^{(1)}(x)=f(x)$，$f^{(k+1)}(x)=f^{(k)}(f(x))(k=1,$

2，…）．因为我们很难从已知条件出发经过计算直接求出 $f(x)$的表达式，故首先构造一个算法求 $f^{(n)}(x)$的表达式．由已知条件有

$$f^{(1)}(x)=f(x)=a_1f(x)+b_1x,\ (a_1=1,\ b_1=0)$$

$$f^{(2)}(x)=-f(x)+6x=a_2f(x)+b_2x\ .$$

$$(a_2=-1,\ b_2=6)$$

由此，猜想 $f^{(n)}(x)$有形式：$f^{(n)}(x)=a_nf(x)+b_nx$．
下面用归纳法构造求 a_n，b_n 的算法：由 $a_{k+1}f(x)+b_{k+1}x=f^{(k+1)}(x)=f^{(k)}(f(x))=a_kf(f(x))+b_kf(x)=a_k\cdot(-f(x)+6x)+b_kf(x)=(-a_k+b_k)f(x)+6a_kx$，得出计算 a_n，b_n 的算法公式：$a_1=1$，$a_2=-1$ 及 $a_{k+1}=-a_k+b_k$ 及 $b_{k+1}=6a_k$．消去 b_k 得 $a_{k+1}=-a_k+6a_{k-1}$，用特征根法易求出 $a_k=\frac{1}{5}[2^n-(-3)^n]$，$b_n=\frac{6}{5}[2^{n-1}-(-3)^{n-1}]$，于是

$$f^{(n)}(x)=\frac{1}{5}[2^n-(-3)^n]f(x)+\frac{6}{5}[2^{n-1}-(-3)^{n-1}]x,$$

即

$$0\leqslant\frac{5f^{(n)}(x)}{3^n}$$

$$=\left[\left(\frac{2}{3}\right)^n-(-1)^n\right]f(x)+2\left[\left(\frac{2}{3}\right)^{n-1}-(-1)^{n-1}\right]x\ .$$

令 n 为偶数且 $n\to+\infty$ 可得 $-f(x)+2x\geqslant0$，即 $f(x)\leqslant2x$；再令 n 为奇数且 $n\to+\infty$ 可得 $f(x)-2x\geqslant0$，即 $f(x)\geqslant2x$；于是得 $f(x)=2x$，且易验证它满足题中条件，这样就通过构造计算 $f^{(n)}(x)$的算法而间接地求出了 $f(x)$的表示式．

2．构造算法逼近原问题的解

例　求代数方程 $f(x)=0$ 的根．

分析：先将 $f(x)=0$ 变成为与之等价的另一方程 $x=$

$g(x)$，任取一个初始值 $x=x_0$，构造一个算法 $x_n=g(x_{n-1})$ $(n=1, 2, \cdots)$，便可逐次算出 x_0，x_1，x_2，…如果 $\lim\limits_{n\to+\infty} x_n = x^*$ 存在，那么在 $x_n=g(x_{n-1})$ 两边令 $n\to+\infty$ 取极限，只要 $g(x)$ 在 x^* 连续就有 $x^*=g(x^*)$，即 x^* 是方程 $x=g(x)$ 的根，从而 x^* 也是 $f(x)=0$ 的根．可见，我们逐步计算出来的 x_0，x_1，x_2，…随着 n 的增加越来越逼近方程 $f(x)=0$ 的根．

例如考察方程 $f(x)=x^2-3=0$．它可变成下列各种形式：

a）$x=x+(x^2-3)$；

b）$x=\dfrac{1}{2}\left(x+\dfrac{3}{x}\right)$；

c）$x=x+\dfrac{1}{4}(x^2-3)$，等等．

对应地可构造下列各种算法公式：

a）$x_n=x_{n-1}+(x_{n-1}^2-3)$；

b）$x_n=\dfrac{1}{2}\left(x_{n-1}+\dfrac{3}{x_{n-1}}\right)$；

c）$x_n=x_{n-1}+\dfrac{1}{4}(x_{n-1}^2-3)$，等等．

我们都从 $x_0=2.0$ 开始，将三个算法的计算结果列表如下：

	(a)	(b)	(c)
x_0	2.0	2.0	2.0
x_1	3.0	1.75	1.75
x_2	9.0	1.732143	1.734375
x_3	87.0	1.732051	1.732361
x_4		1.732051	1.732092

续表

	(a)	(b)	(c)
x_5			1.732056
x_6			1.732051
x_7			1.732051

算法(a)中 $x_n \to +\infty$，算法(b)及(c)中 x_n 都收敛到原方程的解；但很明显，算法(b)的效率比算法(c)好，换句话说，算法(b)比(c)收敛得快。

由于理论和实践的需要，我们设计的算法应该满足一定的要求，对于同一个问题，往往可以有不同的算法，对不同算法的优劣，也应该作出评价．所以，应该讨论算法的评价的问题．评价的标准，通常有以下的三个方面：

(1)构造的算法应该是成功的．所谓成功，指的是按算法得出的数列必须是收敛的，且在收敛的情况下，收敛速度越快的算法越好．有些算法虽然收敛，但收敛的速度太慢(比如说，用电子计算机计算，要达到要求的精度，需要计算几年甚至几十年)，这样的算法在实际使用时与不收敛的算法其实没有什么区别，因而也可以认为收敛太慢的算法实际上是不成功的．至于经过有限步计算即可以得到精确解的算法我们总认为是收敛的，因而是成功的．

(2) 算法所包含的工作量应该越少越好．对于一个规模为 n 的问题(例如，含 n 个变量的方程组，一元 n 次方程，等等)，假设计算所需的工作量为 $S(n)$(常指计算次数或计算时间)，则称 $S(n)$为算法的复杂性函数．因此，算法的工作量问题常称为算法的复杂性问题．

设某个问题的规模为 n，有两种不同的算法，它们计算这

个问题所需工作量(运算次数)分别为:

$$(a)\ S(n)=n^5;\quad (b)\ S(n)=3^n.$$

假定高速电子计算机完成一次运算所需时间为1微秒(0.000001秒). 那么上述规模为 n 的问题所需解题时间 $S(n)$(时间工作量)可列成下表:

时间复杂性函数 $S(n)$	规模				
	5	10	20	40	60
$S(n)=n^5$	0.003125秒	0.1秒	3.2秒	1.7分	13.0分
$S(n)=3^n$	0.000243秒	0.059秒	58分	3855世纪	1.3×10^{13}世纪

从表中可以看出, 当 $S(n)$ 为 n 的多项式时, 工作量随问题规模 n 的增长而增长的速度比较平缓, 但对于 $S(n)$ 为指数函数的情形, 这种增长随着 n 的增大而变得异常激烈. 通常把时间复杂性函数 $S(n)$ 是 n 的多项式的算法称为多项式时间算法, 而把 $S(n)$ 是 n 的指数式的算法称为指数时间算法. 因此, 我们总认为多项式时间算法是"好的"算法.

(3) 算法的稳定性. 由于原始数据是从实际中通过各种仪器测试、记录下来的, 必然存在误差; 另一方面, 用电子计算机进行计算时, 总是用有限小数代替实数, 因而又有"截断误差". 对于一个算法, 如果原始数据的微小变化只引起运算结果的微小变化, 则上述各种误差对计算结果的影响很小, 我们称这样的算法是数值稳定的, 否则就称为数值不稳定的. 对于不稳定的算法, 尽管理论上已经证明经过有限步计算可以求得精确解, 但实际上由于误差的影响和算法的不稳定性, 算出来的结果与精确解相去甚远, 根本不能作为原问题的解答, 因此实际上我们总是要求所构造的算法具有数值稳定性.

综上所述，评价一个算法的好坏，我们总是从收敛性(包括收敛速度)、计算的复杂性、数值稳定性三个方面来考虑，收敛速度快、计算工作量少且数值稳定的算法才是好的算法．

随着电子计算机的发展和普及，计算机已广泛地用于解决生产、管理及日常生活中的许多实际问题．用电子计算机处理这些问题，需要将问题算法化．计算机才能进行操作以至得到问题的结果或解答．问题算法化的过程包括两个步骤：首先是将问题形式化或数学化，即建立适当的数学模型来刻画或拟合要解决的实际问题；其次是寻找合适的算法．这时有两个方面的问题：其一是可计算性，即利用计算机解决问题的可能性的研究．对于具体问题，则应将具体的算法构造出来；其二是关于构造的算法的判定，即研究算法的收敛性、计算复杂性和稳定性．显然．从问题的形式化到构造算法，都离不开数学思想方法的指导．

二、几何定理的机器证明 *

构造算法是一种抽象的构造性思维，它不仅可以用来解决代数中的计算和证明问题，而且可以用来证明几何定理．

几何定理的机器证明只是机器证明或称自动推理的一个方面，由于传统学科采用现代证明手段，它在机器证明的研究中占有重要的地位．我国数学家在这方面做过许多开创性的工作，具有国际领先的优势．

几何问题的证明特别需要智慧与技巧，如何证题在很大程度上依赖于个人的经验和修养，通常是一题一证，而无通法可循．所以，寻求几何证明的通法，一直是许多卓越的科学家的梦想．为此，笛卡儿发明了坐标系，莱布尼兹设想过推理机器，希尔伯特在其名著《几何基础》中给出了一类几何命题的机

械化定理.

电子计算机的出现推动了数学机械化. 50年代, 塔斯基用代数方法证明了初等几何机械化的可能性. 之后, 代数与分析计算问题的机械化的研究取得进展, 但直到1975年仍然没有找到能用计算机判定非平凡几何命题的有效算法.

1976年, 我国著名数学家吴文俊在《中国科学》上发表《初等几何判定问题与机械化证明》一文, 首次给出了用计算机判定非平凡几何定理的有效算法. 用他的方法, 可在微机上很快地证明困难的几何定理. 周咸青发展了吴方法并将其实现为有效的通用程序, 证明了512条非平凡定理, 这是一个重大的突破, 国际上公认这是自动推理领域的一项革命性的工作.

吴方法在本质上仍属代数方法, 其基本思想是以繁复的计算作为代价, 用量的复杂取代质的困难. 所给的"证明"是关于多项式的一大堆计算, 人们难于理解其几何意义, 也难于检验其是否正确.

既然代数方法不能使人完全满意, 于是人们致力于几何定理可读证明自动生成的研究. 这方面的研究开始于60年代, 但30多年来进展甚微, 未能给出哪怕一小类几何定理机器证明的有效算法的程序.

1992年, 我国数学家张景中院士提出"消点法"的思想, 与周咸青、高小山合作, 实现了几何定理可读证明的自动生成. 基于此法编出的算法程序, 已在微机上对数以百计的困难定理完全自动生成了简短的可读证明, 这种方法也能用于立体几何. 张院士还与杨路、高小山、周咸青合作, 把消点法用于非欧几何可读证明的自动生成亦得到成功, 获得了一批非欧几何的新定理, 消点法也可用于几何计算和公式推导.

张景中院士认为他的方法"不以坐标为基础, 也不同于传

统的综合方法，而是以几何不变量为工具，把几何、代数、逻辑和人工智能方法结合起来所形成的开放系统.”它选择几个基本的几何不变量和一套作图规则，并建立与这些不变量和作图规则有关的消点公式．当命题的前提以作图语句输入时，程序可调用适当的消点公式把结论中的约束点逐个消去，最后将所有约束点消完．消点过程记录与消点公式相结合，就是一个具有几何意义的证明．

在多数情况下，消点法也可用纸笔(手工)证明不平凡几何命题．它结束了两千年以来几何证明无定法的局面．当我们不懂方程的时候，算术的四则运算复杂题使我们感到困难，但方程使四则运算复杂题的解法变得简单，使这类问题的处理有一种统一的程式，因而方程可称为解决算术的四则运算复杂题的通法．同样地，消点法是解决几何证明问题的通法．

消点法的成功，使几何定理机器证明的成果在数学教育中的应用有了实现的可能，它被国际同行誉为使计算机能像处理算术那样处理几何的发展道路上的里程碑．

机器证明是算法化思想及其应用的一个重要的方面，随着计算机的不断发展和改进，算法化思想在推进数学乃至整个科学向前发展中将起着越来越大的作用．

第五章 转化思想(上)

本章中我们将讨论第一类转化思想，即特殊化思想、一般化思想和变换思想，着重阐述它的含义、分类及其作用.

§5.1 特殊化思想

什么是特殊化思想？美国著名的数学家和数学教育家波利亚曾经精确地指明了它的内涵："特殊化是从对象的一个给定集合，转而考虑那些包含在这个集合内的较小的集合"，并且指出"我们往往从专门研究对象的全体转变为研究包含在这个全体中的仅仅一个对象"．这就是说，对于某个一般性的数学问题，如果一时难以解决，那么可以先解决它的特殊情况，然后再把解决特殊情况的方法或结论应用或者推广到一般问题上，从而获得一般性问题的解答.

特殊情形相对一般情形而言是比较简单、直观和具体的，因而常常易于找出特殊情形的解答．而且普遍性存在于特殊性之中，一些特殊情形的解答和解答过程常常蕴含着一般问题的解法途径或思路．因此，特殊化思想是探索一般性问题的解题途径的重要思想之一.

将一般化问题转化为特殊问题并不困难，常常可以通过取特例、分类、分步等方法来获得，也就是只须把被研究的对象添上某些限制或适当加强某些条件，求出原集合的一个或若干个子集合即可．由于同一个一般性问题经过不同的特殊化处理

可以得到不同的特殊问题，其中究竟谁最有利于一般化问题的解决呢？这是运用特殊化思想解决问题时遇到的最困难的问题，我们自然希望找出的特殊问题既要本身比较直观、简单，易于找出它的解答，又要求由它的解答易于发现一般性问题的解题途径．可见，运用特殊化思想解决一般性问题的关键在于能否找到一个或几个最佳的特殊问题．

一、寻找特例

1．极端原理

有些现象往往在极端情形下发生，有些性质往往为处于极端状态的元素所具有．为了解决某个问题，先考察某些极端状况，如最大(小)距离、最大(小)面积、最长(短)边、最大(小)值、图形的极限位置或临界位置等．从对极端情形的研究中得到启发，并由此获得解决所研究问题的方法．我们称这种指导解题的思想为极端原理．

作为例子，我们考察第10届IMO试题：求证在四面体$ABCD$中，必有某个顶点，从它出发的三条棱作为三边可以构成一个三角形．我们可以这样来思考：要让结论成立，只需找到一个顶点，从这顶点出发的三条棱中最长的一条小于其余两条之和即可．为此，如图5-1，我们先考察四面体$ABCD$的六条棱中最长的那条棱，不妨设之为AB，于是只要证$BC+BD>AB$或$AC+AD>AB$，从而只须证$BC+BD+AC+AD>2AB$即可．而这不难从$BC+AC>AB$及$BD+AD>AB$得到，于是问题得到解决．

应当注意，运用极端原理解决问题时，寻找的极端元素或极端情形可能不止一个，从而使我们可以获得多种不同的解题途径．如上例中如果我们考察四面体的四个顶点中，由其发出

三条棱长之和最大的那个点，同样可以得到问题的解答。

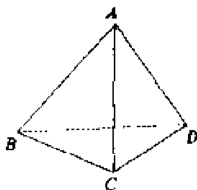


图 5－1

2．不变量思想

许多数学问题中常常含有一些任意性条件，这种任意性条件千变万化，常常使人捉摸不定，难以入手解答。但是透过变化的复杂的现象，常常会发现某种量在问题的变化过程中是不变的，这种不变的量通常称之为不变量。利用某些不变量的特点常常会发现解决问题的途径。这种指导解题的思想就是不变量思想。初等数学问题中遇到的不变量有数值不变量、奇偶性不变量、同余不变量、映射不变量(不动点)、组合不变量等等。运用不变量思想解题的关键是发现或寻求不变量。

例如，我们考察下列问题：已知 $a_1=1$，$a_2=2$，

$$a_{n+2}=\begin{cases}5a_{n+1}-3a_n(a_n\cdot a_{n+1}\text{为偶数}),\\ a_{n+1}-a_n(a_n\cdot a_{n+1}\text{为奇数}),\end{cases}\quad n=1,2,3,\cdots$$

试证：对一切自然数 n，都有 $a_n\neq0$（1988 年全国高中数学联赛第二试第 1 题）。

我们算出此数列的前面几项是 1，2，7，29，22，23，49，26，－17，…．首先易看出其奇偶性规律的不变性：奇、偶、奇；奇、偶、奇；奇、偶、奇；…但仅有这一不变性还不能得出题目结论，因为还必须证明为偶数的项不等于零。其次，注意到奇、偶性实质上是取 2 为模的余数性质。如果我们改为取 3 为模就会发现余数保持 1，2；1，2；1，2；…的规律不变。于是问题归结为证明：$a_{2n-1}\equiv1$，$a_{2n}\equiv2 \pmod 3$。而这一结论不难用数学归纳法证明，从而轻而易举地得到问题的解答。

同样，我们应该注意到，一个复杂问题中的不变量常常不止一个，如果寻找出的不变量不同，那么得到的解题途径也不同．如上例中考察取 4 为模的余数，又可以发现其余数的变化规律为 1，2，3；1，2，3；1，2，3；…，同样可以得到解决问题的正确途径．

3．寻找反例

在数学中，为了证明"某命题不成立"的一类问题，常常运用构造反例的思想方法．所谓数学中的反例，指的是符合某命题的条件，而又不符合该命题的结论的例子．换言之，反例就是指出某命题不成立的例子．一般说来，能说明某命题不成立的反例有很多个，我们只要举出其中一个就够了．

数学中的反例可分为下列几种类型：

（1）关于命题基本形式的反例

数学命题有下列四种基本形式：全称肯定判断（所有 s 都是 p）、全称否定判断（所有 s 都不是 p）、特称肯定判断（存在 s 是 p）、特称否定判断（存在 s 不是 p）．其中全称肯定判断与特称否定判断可以互为反例；全称否定判断与特称肯定判断也可以互为反例．

关于命题基本形式的反例在数学中是很多的．例如，当 $n=0$，1，2，3，4 时，式子$2^{2^n}+1$的值分别是 3，5，17，257，65537，这五个数都是素数．于是，费尔马猜测"对一切非负整数 n，$2^{2^n}+1$ 都是素数"．这是一个全称肯定判断，但后来欧拉举出了特称否定判断的反例："当 $n=5$ 时，$2^{2^5}+1=641\times6700417$ 是一个合数，而不是素数"，从而推翻了费尔马的猜测．

（2）关于充分条件或必要条件的反例

如果"有 A 就有 B"，那么称 A 是 B 的充分条件．这时"没有 A 不一定没有 B"．为此，我们可举出反例"没有 A 却有 B"来说明．这种反例被称为关于充分条件的反例，用它来说明所给条件 A 对结论 B 而言仅仅是充分的．例如，如果两个角是对顶角，那么这两个角相等．但两个角是对顶角仅仅是两角相等的充分条件，因为有的两个角相等但不是对顶角．反例：等腰三角形的两底角相等，而两底角不是对顶角．

如果"没有 A 就没有 B"，那么称 A 是 B 的必要条件．这时"有了 A，不一定有 B"．为此，我们可举反例"有了 A，却没有 B"来说明．这种反例称为关于必要条件的反例，用它来说明 A 对 B 而言仅仅是必要的．例如，$a=b$ 时必有 $a^2=b^2$，但 $a^2=b^2$ 仅仅是 $a=b$ 成立的必要条件，因为存在数 a，b，使 $a^2=b^2$ 但 $a\neq b$．反例：$(-3)^2=3^2$，而 $-3\neq 3$．

(3) 条件变化型反例

当一个数学命题的条件发生变化时，结论不一定正确．为了说明这一点所作出的反例称之为条件变化型反例．例如平行四边形的判定定理之一是：假设四边形 $ABCD$ 满足 1) $AB=CD$；2) $AB/\!/CD$，那么 $ABCD$ 是平行四边形．如果去掉条件 1) 或条件 2)，我们很容易举出等腰梯形这个反例说明，这时结论不再成立．但如果保留条件 1) 而将条件 2) 改为下列条件 2') $\angle B=\angle D$．那么能否得出 $ABCD$ 是平行四边形的结论呢？为了说明这一命题仍不正确，我们可以参看 p.169 图 4-25 的例子．在图 4-25 中，作等边三角形 ABC，在 BC 上取一点 D 使 $BD>DC$，过 D 作 $\angle 2=\angle 1$，并取 $CD=AE$．易证四边形 $ABDE$ 满足条件 1) 和 2')，但它不是平行四边形．

反例在推动数学向前发展和数学教育中有下列一些作用：

(i) 举反例可促进数学新概念、新定理和新理论的形成和发展．例如，公元前500年希帕索斯发现了等腰直角三角形的直角边与斜边的比不再是有理数，这个发现实质上是毕达哥拉斯学派认为的"一切量都可用有理数来表示"的反例．这个反例的出现导致数学史上第一次危机，通过克服第一次危机，使数学大大向前推进了一步．正如美国数学家盖尔鲍姆和奥姆斯特德所指出的"数学由两大类——证明和反例组成，而数学发现也是朝着两个主要的目标——提出证明和构造反例"．对于数学中的重大课题与数学猜想，能举出反例予以推翻，同给出证明予以肯定，是同样重要的．

(ii) 举反例能够澄清数学概念和定理，使人们获得正确的认识．例如我们知道：如果 $f(x)$ 定义在点 x_0 的邻域 $(x_0-\delta,x_0+\delta)(\delta>0)$ 内，且在 $(x_0-\delta,x_0]$ 单调增加而在 $[x_0,x_0+\delta)$ 单调减少，那么 $f(x)$ 在 x_0 取极大值．这是一个正确的命题．然而许多人单凭几何直观认为它的逆命题"如果 $f(x)$ 在 x_0 的某个邻域内有定义且在 x_0 取极大值，那么必存在正数 $\delta>0$，使 $f(x)$ 在 $(x_0-\delta,x_0]$ 单调增加而在 $[x_0,x_0+\delta)$"单调减少"也是正确的．下面的反例推翻了这种看法．

设
$$f(x)=\begin{cases}2-x^2\left(2+\sin\dfrac{\pi}{x}\right), & (x\neq 0);\\ 2, & (x=0).\end{cases}$$

则对任意 $x\in(-\infty,+\infty)$，都有 $f(x)-f(0)=-x^2\cdot\left(2+\sin\dfrac{\pi}{x}\right)\leqslant 0$，故 $f(x)$ 在 $x_0=0$ 处取极大值．然而可以证明对任何 $\delta>0$，$f(x)$ 在 $(0,\delta]$ 不能单调增加，在 $[0,\delta)$ 内不能单调减少．事实上，熟悉微积分的读者容易证明在 $(-\delta,0]$ 和 $[0,\delta)$ 这两个区间内都存在 $f(x)$ 的导数为正的点，也存在 $f(x)$ 的

导数为负的点.

(iii) 在数学中适当地举反例，有助于学生加深对基础知识的理解，提高他们的数学修养，形成严密的逻辑思维的习惯.

在数学教学中，教师应有意识地提出一些思考问题让学生判断是否正确，正确的应给出证明，而错误的则必须举出反例加以否定. 例如在教学复数时，可提出下列命题让学生判断是否正确：1）若 z 为复数，则 $|z|^2=z^2$；2）若 $z_1^2+z_2^2=0$，则 $z_1=z_2=0$；3）若 $a_1+ib_1=a_2+ib_2$，则 $a_1=a_2$，$b_1=b_2$；4）若 $\Delta=b^2-4ac>0$，则方程 $ax^2+bx+c=0$（$a\neq0$）有两个不等实根，等等. 引导学生分别举出反例加以否定后，有助于学生加深对复数概念及其基本性质的理解和掌握.

二、分类思想

当被研究的数学问题出现多种不同的情况时，常常按可能出现的各种情况分别进行讨论和解答，得出各种情形下相应的结论，综合起来就获得原问题的解答，这种指导解决问题的思想称之为分类思想.

因为分类思想是将母项分拆成许多子项，母项与子项的关系就类似于整体与局部的关系，而整体与局部在一定的条件下可以互相转化，整体包含局部，局部性质反映了整体的属性，所以，分类思想在解题中的作用就体现在：化一为多、化复杂为简单、化多为少、化无限为有限. 换言之，要选择一个适当的分类标准，将原来比较复杂的问题化为若干子问题，而选择分类标准的关键在于分类后得到的各个子问题应当是熟悉的或易于解决的问题，从而能通过对各个子问题的解决而最终解决原问题.

应用分类思想解题还必须遵循以下原则：

1. 分类必须包含原题目中所有可能出现的各种情况，没有遗漏；

2. 任何两类之间互相排斥，没有重迭；

3. 每次分类必须使用同一标准；

4. 各类情形都应比原问题易于解决.

例如，我们考察下列问题：平面上给出了5个点，其中任意3点不共线，证明：5点中必有4点是一个凸四边形的四个顶点.

我们可按5点的凸包分为三种情形：(1) 5点的凸包为凸五边形；(2) 五点的凸包为凸四边形；(3) 五点的凸包为$\triangle ABC$，其余两点D，E在$\triangle ABC$内. 对于情形(1)和(2)结论显然成立，对于情形(3)，因为任意三点不共线，直线DE必与$\triangle ABC$的两边相交而与第三边不相交，设直线DE与线段BC不相交，则四边形$EDBC$为凸四边形(如图5－2). 于是，通过分类讨论，我们轻而易举地得到问题的解答.

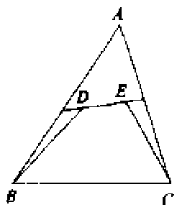


图5－2

应当注意，对于同一个较复杂的问题，如果取用的分类标准不同，那么解决问题的繁简程度往往差别很大. 如上例中，如果我们先以其中三点为顶点作一个三角形，再按其余两点关于这个三角形的位置的不同情况分类进行讨论，就显得异常的复杂，大大增加了解题的难度. 因此，进行分类讨论之前，应当对题目先作一番深入的分析，并将几种分类标准作适当比较，选择恰当的分类标准，尽量简化分类过繁的现象. 其次，对于某些数学问题，如果先利用已知条件，限制题目中涉及的量的范围，再进行分类讨论，常常能使问题简化.

例 已知 k 为正整数，若方程 $kx^2-2(1-2k)x+4k-7=0$ 至少有一个整数根，求 k 的值．

分析：若按常规思路，先解出 $x=\frac{1-2k\pm\sqrt{1+3k}}{k}$，然后就 k 的取值进行分类讨论．显然，这种讨论是十分繁杂的．但如果注意到"k 是正整数"这一已知条件，由原方程得 $k=\frac{2x+7}{(x+2)^2}\geqslant 1$（$x\neq-2$），解得 $-3\leqslant x\leqslant 1$．于是仅对 $x=-3$，-1，0，1 四种情形分类讨论，就很容易得出问题的答案为 $k=1$ 或 $k=5$．

最后，对于某些含有分类讨论因素的问题，如果能根据题目条件中的特征，灵活采用一定的解题策略，有时可避免进行复杂的分类讨论．

例 求同时满足下列两个条件的所有复数：（A）$z+\frac{10}{z}$ 是实数，且 $1<z+\frac{10}{z}\leqslant 6$；（B）$z$ 的实部和虚部都是整数．

分析：按照常规思路，设 $z=x+iy$（x，$y\in R$）代入，再对 x，y 的取值情形分类进行讨论，虽然能求出问题的正确解答，但解答过程较繁．如果注意到 $z+\frac{10}{z}$ 为实数，设 $z+\frac{10}{z}=t$，则 $1<t\leqslant 6$，且易求出 $z=\frac{t}{2}\pm\frac{1}{2}\sqrt{40-t^2}\,i$，再由 z 的虚部 $\frac{1}{2}\sqrt{40-t^2}$ 为整数及 $1<t\leqslant 6$ 得出 $t=2$ 或 6，于是得到 $z=1\pm 3i$ 或 $3\pm i$．

总之，应用分类思想指导解题应做到：定标准（即选定分类的标准）、定类别（即按选定的标准将问题正确地划分为各个类别）、定顺序（即对划分的各类别有序地分层次进行讨论）、

定结论(即综合各类情形的结果，确定该问题的最后结论)，以达到化整为零、化难为易、化繁为简，直到最终解决问题的目的：

三、分步思想

对于一个复杂的数学问题，我们一下子解决不了的时候，就必须考虑分几步走，也就是将原问题分解成一串彼此关联的小问题，在这一串问题中，后一个问题的解决依赖于前一问题的结果，当最后一个问题解出时，原问题也就得到了解答．这种解决问题的思想称为分步思想．分类和分步都是对原问题进行分解，它们的区别在于运用分类思想解题时，所分成的各类情形是彼此独立的，其中任何一类情形的解决不依赖另一种情形，而分步中的一串问题并不互相独立，而是后面的与前面的紧密相关．

在数学研究和解题中常用的递进、递推、逼近等思想方法都属于分步思想的范畴．

1．递进思想

如果我们遇到一个无法直接从已知条件导出结论的问题，可以尝试在已知条件(起点)和结论(终点)之间设立若干个中途点，这就是说，我们可以尝试建立一连串的小问题，每解决一个小问题就使我们的解题过程向前推进了一步，这样逐步推进．通过一连串小问题的解决使我们从起点出发，通过了一个个中途点最终到达终点，将原问题全部解决．这种解决数学问题的思想称为递进思想，又可称为中途点思想．

例 已知一个四位数是完全平方数，它的前两位数字相同，后两位数字也相同，求这个四位数．

分析：设所求四位数为 N，它的前两位数字为 x，后两

位数字为 y，则有 $N=1100x+11y$，直接从已知条件（N 是完全平方数），求出 x，y 较困难，我们分两步走。

第一步求出 N 为完全平方数时，x，y 应满足什么条件？

第二步根据第一步中求出的条件来确定 x，y 应取什么值才能使 N 为完全平方数。

第一步的解答为：由 $N=11(100x+y)=11(11\times 9x+x+y)$ 为完全平方数得 $x+y$ 是 11 的倍数，再由 $1\leqslant x+y\leqslant 18$ 得 $x+y=11$。（中途点到达）

第二步解答为：由 $N=11(11\times 9x+x+y)=11(11\times 9x+11)=11^2(9x+1)$ 为完全平方数知 $9x+1$ 必须为完全平方数。以 $x=1$，2，3，…，9 代入知仅当 $x=7$ 时，$9x+1=8^2$ 为完全平方数，所以仅当 $x=7$，$y=4$ 时 N 为完全平方数，于是得所求四位数为 7744（终点到达）。

对于一个具体的问题来说，困难在于应怎样分步递进（即应怎样确定中途点）。一般说来，针对具体的问题要进行具体分析。通常的方法是：一方面从已知条件出发，盯住结论逐步往下推。另一方面从结论出发，注意已知条件往上找，而所谓顺推倒溯，也就是综合与分析。顺推的好处是每一步都在作正确推理，所以能扩大对问题的了解，即使所得到的中途点并不一定正确，也会获得一些失败的教训，并从中得到启发，使我们能找到正确的解题途径（即找到能通向终点的中途点）。倒溯的好处是解题目的明确，常常可使解题少走弯路，找到一条通向终点的捷径。

例 设 $\{a_n\}$ 是满足

$$1=a_0\leqslant a_1\leqslant a_2\leqslant\cdots\leqslant a_n\leqslant\cdots \quad ①$$

的实数列，而序列 $\{b_k\}$ 由下式定义：

$$b_n=\sum_{k=1}^{n}\left(1-\frac{a_{k-1}}{a_k}\right)\frac{1}{\sqrt{a_k}}. \qquad ②$$

证明：对所有 $n=1, 2, 3, \cdots$有 $0\leqslant b_n<2$ 成立.

分析：由已知条件①知$\left(1-\frac{a_{k-1}}{a_k}\right)\frac{1}{\sqrt{a_k}}\geqslant 0$，故有 $b_n\geqslant 0$．为了推出 $b_n<2$，必须将 b_n 的表达式②变形，使和式便于估计．联想到

$$\sum_{k=1}^{n}\left(\frac{1}{a_{k-1}}-\frac{1}{a_k}\right)=\frac{1}{a_0}-\frac{1}{a_n}=1-\frac{1}{a_n}<1;$$

自然将②改写为

$$b_n=\sum_{k=1}^{n}\left(\frac{1}{a_{k-1}}-\frac{1}{a_k}\right)\frac{a_{k-1}}{\sqrt{a_k}}. \qquad ③$$

如果有$\frac{a_{k-1}}{\sqrt{a_k}}\leqslant 2$，那么结论便可得证．但满足题目条件①的数列不保证满足$\frac{a_{k-1}}{\sqrt{a_k}}\leqslant 2$．例如 $a_k=k$ 时，$\frac{a_{k-1}}{\sqrt{a_k}}\to+\infty(k\to+\infty)$．虽然失败了，但分析失败的原因是$\frac{a_{k-1}}{\sqrt{a_k}}$中分子的次数为1，高于分母的次数$\frac{1}{2}$，再考虑将分子分母都变成$\frac{1}{2}$次，为此，将③再变形为

$$\begin{aligned}b_n&=\sum_{k=1}^{n}\left(\frac{1}{\sqrt{a_{k-1}}}+\frac{1}{\sqrt{a_k}}\right)\left(\frac{1}{\sqrt{a_{k-1}}}-\frac{1}{\sqrt{a_k}}\right)\frac{a_{k-1}}{\sqrt{a_k}}\\&=\sum_{k=1}^{n}\frac{\sqrt{a_{k-1}}}{\sqrt{a_k}}\left(1+\frac{\sqrt{a_{k-1}}}{\sqrt{a_k}}\right)\left(\frac{1}{\sqrt{a_{k-1}}}-\frac{1}{\sqrt{a_k}}\right). \qquad ④\end{aligned}$$

又由已知条件①有

$$\frac{\sqrt{a_{k-1}}}{\sqrt{a_k}}\left(1+\frac{\sqrt{a_{k-1}}}{\sqrt{a_k}}\right)\leqslant 2,\quad \frac{1}{\sqrt{a_{k-1}}}-\frac{1}{\sqrt{a_k}}\geqslant 0, \qquad ⑤$$

于是得到

$$b_n \leqslant 2\sum_{k=1}^{n}\left(\frac{1}{\sqrt{a_{k-1}}}-\frac{1}{\sqrt{a_k}}\right)=2\left(\frac{1}{\sqrt{a_0}}-\frac{1}{\sqrt{a_n}}\right)$$

$$=2\left(1-\frac{1}{\sqrt{a_n}}\right)<2.$$

从以上"顺推倒溯"的过程中得到一条从已知到结论的通路：

已知→③→④→⑤→结论.

其中③、④、⑤便是通过分析与综合找出的中途点.

总之，解题中运用递进的思想，恰当地选择中途点，就能化复杂问题为简单问题，化长途为短途，中途点就好像登山时的一级级台阶，沿着台阶一级又一级向上攀登，则一定能达到登上高峰的目的.

2. 递推思想

对于已知条件和结论都与自然数相关的问题，常常可先确定一个或几个初始值，然后一个一个地进行计算、观察、分析和比较，并概括其规律，从而建立某种递推关系，最后利用递推关系解题. 这种用于解决问题的思想称为递推思想.

例 把一个圆分成 n 个不相等的扇形，依次记为 S_1，S_2，…，S_n（$n\geqslant 2$）. 每个扇形可用红、黄、蓝三色之一着色，但要求相邻的两个扇形不同色，那么有多少种着色方法？

分析：设着色方法总数为 x_n，易知 $x_2=3\times 2=6$. 下面考察 x_n 与 x_{n-1} 的关系. 首先 S_1 有 3 种染法，则 S_2 只能染 S_1 以外的两色，有 2 种方法，…，S_n 只能染 S_{n-1} 外的两色，有 2 种方法，于是共有 $3\times 2^{n-1}$ 种方法. 但这样仅保证了 S_i 与 S_{i-1}（$2\leqslant i\leqslant n$）不同色，不能保证 S_n 与 S_1 不同色. 故有两种可能：若 S_n 与 S_1 不同色，则有 x_n 种方法；若 S_n 与 S_1 同色，则可将 S_n 与 S_1 合在一起看成一个扇形，从而有 x_{n-1}

种方法，于是得到递推关系：$x_n+x_{n-1}=3\times2^{n-1}$．解这个递推关系得$x_n=2^n+(-1)^n\cdot2$，此即为所求的着色方法数．

从上例可以看出，利用递推思想指导解题的步骤是：第一步求初始值；第二步建立递推关系；第三步利用递推关系解题．上述三步中最关键的一步是建立递推关系．如果递推关系是由不完全归纳法或由观察猜测的，那么递推关系的正确性还必须进行严格的证明．递推关系常见的形式除了等式递推关系外，还有不等式递推关系、同余递推关系、整除递推关系等等．需要建立什么形式的递推关系，常常可由问题的性质来确定．其次，求初始值是很重要的，因为没有初始值，递推关系就因缺乏基础而成为空中楼阁．初始值应求几个，应由递推关系来确定．例如递推关系为$x_n=f(x_{n-1},x_{n-2},\cdots,x_{n-k})(n\geqslant k+1)$时，初始值必须求出$k$个：$x_1$，$x_2$，…，$x_k$，才可能利用上述递推关系进行下一步的计算或推理．然而我们常常在建立递推关系前求初始值，这样做的好处是从简单、特殊情形做起有助于弄懂题意，同时从中常常可以发现相邻项之间的联系，从而有助于递推关系的建立．不过建立好递推关系后应回过头来检查一下，看求出的初始值的个数是否符合要求，多了则可以去掉，不够时则应补求几个．最后，求解递推关系或利用递推关系进行推理，除了几种特殊的递推关系(例如常系数线性递推关系)外，并无一个固定的模式，常常需要使用一系列数学方法和技巧，例如利用迭代、叠加、累乘等手段来解决问题．比如上例中递推关系$x_n+x_{n-1}=3\times2^{n-1}$，可通过令$y_n=\frac{x_n}{2^n}$及$z_n=y_n-1$化为$z_n=-\frac{1}{2}z_{n-1}$解出，也可以直接迭代解出：$x_n=3\times2^{n-1}-x_{n-1}=3\times2^{n-1}-3\times2^{n-2}+x_{n-2}-\cdots$
$=3\times2^{n-1}-3\times2^{n-2}+\cdots+(-1)^{n-3}3\times2^2+(-1)^{n-2}3\times2=$

$(-1)^{n-2}\cdot 3\times 2\cdot\dfrac{1-(-2)^{n-1}}{1-(-2)}=(-1)^n\cdot 2+2^n$. 此外，有些问题仅仅只有一种递推关系，还很难解决问题．此时，还必须进一步探求新的递推关系．

例 设 $a_1=1$，$a_2=-1$，$a_n=-a_{n-1}-2a_{n-2}(n\geqslant 3)$．证明：当 $n\geqslant 2$ 时，$2^{n+1}-7a_{n-1}^2$ 是一个完全平方数．

分析：我们只须证明当 $n\geqslant 2$ 时，存在整数 $p\geqslant 0$ 及整数 q 使下列新的递推关系成立：$2^{n+1}-7a_{n-1}^2=(pa_n+qa_{n-1})^2$，而由 $a_1=1$，$a_2=-1$，$a_3=-1$，$a_4=3$ 代入得 $p=2$、$q=2$，故只须证明 $2^{n+1}-7a_{n-1}^2=(2a_n+a_{n-1})^2$，而这不难利用数学归纳法证明．于是，通过建立上述新的递推关系，使问题立即得到解决．

3、逼近思想

逼近思想可以一直追溯到我国古代数学家刘徽用“割圆术”来求圆周率 π 的近似值．所谓“割圆术”就是用一系列圆内接正多边形的周长和面积来逼近圆的周长和面积．当圆内接正多边形的边数逐渐增加，其周长和面积就越来越接近圆的周长和面积．用数学语言来讲就是圆的周长和面积分别是圆内接正多边形的周长和面积当边数无限增多时的极限值．极限思想就是一种最基本的逼近思想，它在现代数学分析中有着重要的、不可替代的作用．在实数理论中用有理数逼近无理数；在微分学中用平均变化率逼近瞬时变化率；在积分学中用有限和逼近无限和；在级数理论中用多项式逼近函数等等都是极限思想的具体运用．又如在求微分方程的解时，常常用差分法或有限元法将微分方程离散化得到一个线性方程组，求出线性方程组的解就得到微分方程的近似解．且当划分越来越细时，所得线性方程组的解就越来越逼近微分方程的解．因此，离散化思想也是

一种重要的逼近思想．当然，逼近并非只是求极限和用离散变量代替连续变量进行近似计算．一般说来，在解决数学问题的过程中，常常一开始放宽条件、降低要求得出问题的一个初步解答，并以此结果为基础，逐步加强要求，使之逐步符合其余的、尚未满足的条件，直到满足所有要求的条件为止，从而得到原问题的解答．这种用来指导解题的数学思想就是逼近思想．由于应用逼近思想解题，具有分散难点，逐步突破难点的功能，因此，它在解题中起到非常重要的作用．在解数学题时，常常用到的逐步调整思想和逐步淘汰思想都属于逼近思想的范畴．

(1) 逐步调整思想

逐步调整思想就是根据问题的特点，首先确定一个初始状态，然后从对相关变量的调整(变换)入手逐步安排试验，每进行一次调整后，就对试验结果进行分析比较，找出变量变化规律，朝有利于解决问题的方向再安排下一次试验，从而最终得到所需要结论的一种思想方法．

利用逐步调整思想解题的关键是设计调整方向并确定具体的调整过程．

利用调整思想解题有下列三种基本类型．

a) 局部扰动逼近

对于一个含多个变量(因素)的问题，可以先假设其中一部分变量不动，而让其余变量变化，求出问题的局部解，再回过头来调整那些暂时固定不动的变量，从而求出问题的整体解．这种从局部解，通过扰动(即调整)逼近整体解的思想方法常常称为局部扰动逼近法．

例 将正方形 $ABCD$ 分割为 n^2 个相等的小方格(n 为自然数)，把相对顶点 A，C 染成红色，把 B，D 染成蓝色，其

他交点任意染成红蓝两种颜色中的一种颜色．证明：恰有三个顶点同色的小方格数目必是偶数．

分析：$n=1$ 时，恰有三个顶点同色的小方格个数为 0，它是一个偶数，命题成立．

下设 $n\geqslant 2$．首先考察特殊情形(初始状态)：除 B，D 外，其余各交点皆染红色，这时恰有三个顶点同色的小方格个数为 2，命题结论成立．

其次，再逐步进行调整，每一步调整都是将一个红色交点变为蓝色交点．则以此点为顶点的小方格要么由三顶点同色变为非三顶点同色，要么由非三顶点同色变为三顶点同色．可见，每调整一次，不改变恰有三顶点同色的小方格个数的奇偶性，而任何染色方式都可由初始状态经过有限步上述调整而得到．所以，在一般情形下，恰有三顶点同色的小方格个数与 2 的奇偶性相同，即为偶数．

b）逐步构造逼近

为了证明存在某种对象同时满足若干条件，我们可以先构造一个对象使之满足其中一部分条件，再对构造出的对象进行调整，使之逐步满足其余的条件，直到构造出的对象满足全部所要求的条件为止．这种解决问题的思想方法称之逐步构造逼近法．

例 平面上任给 100 个点，证明：在此平面内总存在若干个两两无公共点的圆面盖住这 100 个点，且这些圆面的直径之和不大于 100，而每两个圆面的距离大于 1．

分析：1°．首先，以给定的100个点为中心，$\frac{1}{2}$ 为半径作圆．这 100 个圆的直径之和为 100．

2°．第一次调整：使盖住100个点的圆面两两无公共点且

直径之和不增加．办法是：如果有两圆 O_1 与 O_2 相交或外切，那么可以在线段 O_1O_2 上取某点为中心作圆 O_3 使⊙O_1 和⊙O_2 都内切于⊙O_3．显然⊙O_3 的直径小于或等于⊙O_1 与⊙O_2 的直径之和且⊙O_3 内含的已知点到⊙O_3 的周界的距离不小于$\frac{1}{2}$．这时，我们用⊙O_3 代替⊙O_1 和⊙O_2．在有两圆面相交或外切的情形下，一直这样做下去，直到各圆面两两无公共点为止．于是，我们作出了 k $(1\leqslant k\leqslant 100)$ 个直径之和不大于 100 的且两两无公共点的圆面，它们不仅盖住了 100 个已知点，而且每个已知点到盖住它的那个圆面的周界的距离不小于$\frac{1}{2}$．设这 k 个圆两两之间的距离中最小的为 d．则 $d>0$．

3°．若 $d>1$，则所作的 k 个圆满足题目全部条件．若 $0<d\leqslant 1$，则作第二次调整：适当减少这个圆的半径，使能满足全部题设条件．事实上，只要每个圆的半径减少$\frac{1}{2}-\frac{d}{3}>0$，则得到的 k 个圆面仍盖住了已知 100 个点，且直径之和不大于 100．而任意两圆面的距离之和 $2\cdot\left(\frac{1}{2}-\frac{d}{3}\right)+d-1+\frac{d}{3}>1$．

c）"磨光"式单调逼近

所谓"磨光"，就是逐步消灭问题状态间的差异，使之达到平衡或均匀状态．它通常的作法是：根据"磨光"目标，进行适当调整，减少状态间的差异，而越来越逼近目标．我们称这种解题的思想方法为"磨光"式单调逼近法．

例 已知 $0<\alpha_i<\pi$ $(i=1, 2, \cdots, n)$，且 $\alpha_1+\alpha_2+\cdots+\alpha_n=A$．求证：$\sin\alpha_1+\sin\alpha_2+\cdots+\sin\alpha_n\leqslant n\sin\left(\frac{A}{n}\right)$．

分析：记 $S(\alpha_1,\alpha_2,\cdots,\alpha_n)=\sin\alpha_1+\sin\alpha_2+\cdots+\sin\alpha_n$，则

要证不等式为 $S(\alpha_1,\alpha_2,\cdots,\alpha_n)\leqslant S\left(\frac{A}{n},\frac{A}{n},\cdots,\frac{A}{n}\right)$. 因此只要我们将 α_1、α_2、…、α_n 逐个调整到$\frac{A}{n}$，而每次调整不会使和 $S(\alpha_1,\cdots,\alpha_n)$增加，则不等式即可获证.

不妨设 $\alpha_1\leqslant\alpha_2\leqslant\cdots\leqslant\alpha_n$. 因为 $\alpha_1+\alpha_2+\cdots+\alpha_n=A$，于是 $\alpha_1\leqslant\frac{A}{n}$，$\alpha_n\geqslant\frac{A}{n}$，将 α_1、α_n 分别调整为 $\alpha'_1=\frac{A}{n}$，$\alpha'_n=\alpha_1+\alpha_n-\frac{A}{n}$，则 $\alpha'_1+\alpha_2+\alpha_3+\cdots+\alpha'_n=A$，而 $S(\alpha_1,\alpha_2,\alpha_3,\cdots,\alpha_n)-S(\alpha'_1,\alpha_2,\alpha_3,\cdots,\alpha'_n)=\sin\alpha_1+\sin\alpha_n-\left[\sin\frac{A}{n}+\sin\left(\alpha_1+\alpha_n-\frac{A}{n}\right)\right]=4\sin\frac{\alpha_1+\alpha_n}{2}\sin\frac{\frac{A}{n}-\alpha_1}{2}\sin\frac{\frac{A}{n}-\alpha_n}{2}\leqslant 0$, 所以 $S(\alpha_1,\alpha_2,\alpha_3,\cdots,\alpha_n)\leqslant S\left(\frac{A}{n},\alpha_2,\alpha_3,\cdots,\alpha'_n\right)$，其中 $\alpha_2+\alpha_3+\cdots+\alpha'_n=\frac{n-1}{n}A$，再对 α_2，α_3，…，α'_n 作如上调整，至多调整 $n-1$ 次，可使每个 α_i 都变成$\frac{A}{n}$，且有 $S(\alpha_1,\cdots,\alpha_n)\leqslant S\left(\frac{A}{n},\frac{A}{n},\cdots,\frac{A}{n}\right)$，从而得到要证的不等式.

总之，从上述一些例题可以看出，用逐步调整思想解题的主要技巧在于调整试验的安排，巧妙的试验往往能使困难的问题迅速得到解决. 由于调整的形式多种多样，因此，用逐步调整思想解决的问题类型也不限于上述三种. 此外，在用逐步调整思想解题时，常常还要与其他思想方法(如排序思想、极端思想等)密切配合，才能使问题更易获得解决.

(2) 逐步淘汰思想

当我们需要找出满足一定条件的对象时，我们可用给定的条件(“筛子”) 为依据，对所研究的对象进行考察，把符合条

件的对象选出来，而把不符合条件的对象淘汰掉，最后得到所需要的结果．这种用来指导解题的思想称为逐步淘汰思想，又称筛法思想．用筛法思想解题的著名例子可追溯到公元前3世纪希腊数学家爱拉多斯尼对自然数 N 以内的素数的探求：先将1删去，再将大于2的2的倍数删去，然后将大于3的3的倍数删去，接着又把大于5的5的倍数删去，…，如此下去，直到将1及所有合数删去为止．剩下的便是要求的素数．

用逐步淘汰思想解题的关键是恰当地构造“筛子”，也就是确定淘汰的标准．由于所遇到的问题不同，构造“筛子”的方法也因题而异，应当针对具体问题作具体分析．

利用逐步淘汰思想解题有下列四种基本类型．

a）直接淘汰法

它是对所研究的对象直接进行筛选，找出符合条件的对象．

例 将1000表示成若干个正整数之和，使得这些正整数之积为最大．

分析：设将1000表示成 $x_1+x_2+\cdots+x_n=1000$（x_1，x_2，…，x_n 均为正整数)后，使乘积 $p=x_1x_2\cdots x_n$ 取最大值．且不妨设 $x_1\leqslant x_2\leqslant\cdots\leqslant x_n$，于是

（1）$x_1\neq 1$，否则 $1000=x_2+x_3+\cdots+(x_n+1)$，而 $p'=x_2x_3\cdots(x_n+1)>1\cdot x_2\cdots x_n=p$．

（2）$x_n\leqslant 4$．这是因为若 $x_n\geqslant 5$，将 x_n 换为两个数 x_n-2 和2，则 $1000=x_1+x_2+\cdots+x_{n-1}+2+(x_n-2)$，而 $p'=x_1x_2\cdots x_{n-1}\cdot 2\cdot(x_n-2)>x_1\cdot x_2\cdots x_n=p$．

（3）x_1，…，x_n 中可以不出现4．这时因为每个4可用两个2代替，其和与积均不变．

（4）x_1，x_2，…，x_n 中不能有3个或3个以上的2．这

是因为3个2可用2个3代替，其和不变：$2+2+2=3+3$，而乘积增大：$3\times3>2\times2\times2$．

通过以上筛选知，x_1，…，x_n 只能为2或3，且2的个数不多于2个．于是，将1000表为 $332\times3+2\times2$ 时，其乘积取最大值 $P=3^{332}\cdot2^2$．

b）间接筛选法

它不是直接对研究的对象进行筛选，而是对那些与所研究的对象密切相关的某些其他对象进行筛选，通过对这些"相关的"对象进行研究和筛选，间接地筛选出符合条件的对象．

例 黑板上写着前 n 个自然数：1，2，…，n．擦去其中一个自然数后，剩下的各数的平均值是33，且知道擦去的数既不是1，也不是 n，问擦去的数是什么？

分析：本题的困难在于既不知道 n 的值，又不知道擦去的数为何数．先确定 n 的值是关键．只要将 n 的值筛选出来，擦去的数就容易找到．先考察两种极端情形：若擦去的数为1，则剩下的数的平均值为 $\frac{1}{n-1}(2+3+\cdots+n)=\frac{1}{2}(n+2)$；若擦去的数为 n，则剩下的数的平均值为 $\frac{1}{n-1}[1+2+\cdots+(n-1)]=\frac{1}{2}n$．而已知擦去的数既不是1，又不是 n，且剩下的数的平均值为33，故应有 $\frac{n}{2}<33<\frac{1}{2}(n+2)$，即 $64<n<66$，所以 $n=65$．最后设擦去的数为 x，则 $\frac{1}{64}(1+2+\cdots+65-x)=33$，解得 $x=33$，即擦去的数为33．

c）比较筛选法

在解决实际问题时，常常要进行一系列试验，我们希望从中找出最好的试验结果来，常用的作法是开始进行一个或几个

试验，以后每进行一次试验，就将所得结果与前面相邻的一个或几个试验结果进行比较，将最差的一个试验结果淘汰，其余的结果留下，再安排下一次试验，直到得到比较满意的结果为止．这时，因为什么是最好的、希望的试验结果，我们事先并不知道．因此，每进行一次试验后，不可能通过与最好结果比较来进行筛选，而只能将相邻的试验结果比较后进行筛选，故我们称这种解决问题的思想方法为比较筛选法．

在优选法理论以及非线性方程求根理论中所用的对分法、0.618 法等等都是这种思想方法的具体运用．

例 求 $f(x)=x^3-2x-5=0$ 在$[2,+\infty)$内的实根，试求这个实根的近似值．

分析：因$f(2)=-1<0$，$f(3)=16>0$，知 $f(x)=0$ 在区间$(2,3)$内有一个实根．又 $g(x)=x^3-2x=x^2\cdot(x-2)$在$[2,+\infty)$内单调增加，所以 $f(x)=g(x)-5$ 在$[2,+\infty)$内单调增加．故 $f(x)=0$ 在$[2,+\infty)$内只有惟一实根 x^* 属于区间$(2,3)$．

取 $x_1=\frac{1}{2}(2+3)=2.5$，由 $f(2.5)\cdot f(2)<0$，知所求实根 x^* 必在$(2,2.5)$内（这时将 3 筛去）；

取 $x_2=\frac{1}{2}(2+2.5)=2.25$，由 $f(2)\cdot f(2.25)<0$，知所求实根 x^* 必在$(2,2.25)$内（这时将 2.5 筛去）；

取 $x_3=\frac{1}{2}(2+2.25)=2.125$，由 $f(2)\cdot f(2.125)<0$，知所求实根 x^* 必在$(2,2.125)$内（这时将 2.25 筛去）；

取 $x_4=\frac{1}{2}(2+2.125)=2.0625$，由 $f(2.125)\cdot f(2.0625)<0$ 知所求实根 x^* 在$(2.0625,2.125)$内（这时将 2 筛去）；

取 $x_5=\frac{1}{2}(2.0625+2.125)=2.09375$，由 $f(2.09375)\cdot f(2.125)<0$ 知所求实根 x^* 在(2.09375，2.125)内(这时将 2.0625 筛去)；

取 $x_6=\frac{1}{2}(2.09375+2.125)=2.109375$，由 $f(2.09375)\cdot f(2.109375)<0$ 知所求实根 x^* 在(2.09375，2.109375)内(这时将 2.125 筛去)；

取 $x_7=\frac{1}{2}(2.09375+2.109375)=2.1015625$．于是 $|x^*-x_7|<\frac{1}{2}(2.109375-2.09375)=0.0078125$．所以 x^* 的近似值为 2.101，误差小于 0.008．

d）分类筛选法

当研究对象存在的范围较广，不易进行整体比较筛选，这时，我们可以采取分类的办法，逐一进行筛选，使问题变得简单而易于解决．

例 求出所有自然数 n，使一个整数 n、$n+8$，$n+16$ 都是素数．

分析：将 n 分为三类：$3k$，$3k+1$，$3k+2$．当 $n=3k$ 时，只有 $k=1$ 时 $n=3$ 为素数，这时 $n+8=11$ 及 $n+16=19$ 皆为素数；当 $n=3k+1$ 时，$n+8=3k+9=3(k+3)$不是素数；当 $n=3k+2$ 时，$n+16=3k+18=3(k+6)$不是素数．所以，满足题目要求的自然数只有一个 3．

由前面的例子可以看出，对于那些不易(或无法)直接找出结论的问题，如果能够根据题设条件找出各种可能的情形的话，那么一般可用筛法对各种情形进行筛选，而且往往能奏效．筛法主要用于处理数论和组合问题．随着数学日新月异地

发展，爱拉多斯尼创立的筛法已得到重大的发展和改进，筛法已成为现代解析数论中最重要的思想方法之一．特别值得指出的是我国著名数学家陈景润在 1973 年的《中国科学》上发表了用筛法证明“任何一个充分大的偶数可表示成一个素数与另一个至多有两个素因子积的整数之和”(这一结论简称为(1+2))的著名论文后，在国际上引起了巨大的轰动．由于（1+2）这一结论的重要性，人们曾先后对它给出至少五个简化证明，其中最简单的证明是我国三位著名数学家王元、丁夏畦和潘承洞给出的．陈景润的这一结果在哥德巴赫猜想领域至今保持着领先地位．此外，在对命题(1+2)的证明中，陈景润还创造并使用了一种新的加权筛法，后来在数论研究中得到广泛地应用．他所采用的把估计某种集合中元素个数的问题转化为计算另一集合中元素个数的问题的思想，被国际数学界称为转换原理，在数学中有广泛地应用．

§5.2 一般化思想

当我们遇到某些特殊问题感到很难解决时，不妨适当放宽条件或改变一些条件的限制，把待处理的特殊问题放在一个更为广泛、更为一般的问题中加以研究，先解决一般情形，再把解决一般情形的技巧、方法或结果应用到特殊问题上，最后获得特殊问题的解决．这种用来指导解决问题的思想称之为一般化思想．

由于从一般性问题入手可使我们的视野更为广阔，避免在枝节问题上纠缠，容易触及问题的本质．同时，由于限制条件减少，涉及范围增大，更容易引起联想，发现各种条件与结论之间的内在联系．此外，由于人们对一些一般性事物(如数学

模型、公式等)较为熟悉，反而对特殊情形不甚了解．因此，从一般到特殊是人们认识事物的另一个重要侧面．它与从特殊到一般是相辅相成、和谐统一的两个方面．

一、模型化思想

所谓数学模型，常常指那些直接描述现实原型的数学关系结构．广义地说，凡是从现实原型抽象概括出来的一切数学概念、数学公式、函数、方程式、定理和理论体系等等都可称为数学模型．

数学中的模型化思想是指通过建立数学模型来研究现实原型的内在性质、特征及其规律的一种思想方法．

实际问题千差万别，需借助于数学模型来加以研究的问题不计其数，但这些数学模型大体可分为四类．

第一类是确定性数学模型．这类模型反映的现实原型具有确定性和固定性，也就是对象之间的相互联系和因果关系是必然的．这种数学模型的数学形式可以是各种各样的函数、方程式(包括代数方程、微分方程、差分方程、积分方程等等)、关系式（包括逻辑关系式)、网络图等等．由此可见，对确定性模型使用的是经典的数学思想方法．

第二类是随机数学模型．这类模型是用来处理大数现象的，也就是所研究的对象的变化发展往往具有多种不同的可能性，但究竟出现哪种结果，却带有偶然性或随机性，且随着事件的大量出现而呈现一定的统计规律．对这类模型使用的数学思想方法主要是概率论和数理统计的思想方法，即随机思想方法．

第三类是模糊性数学模型．由于客观现实中许多对象及其关系具有模糊性(如“高个子”、“比较年轻”、“能力较强”等

等)，用确定性、随机性数学模型来处理这类问题不十分有效，这就需要建立模糊性的数学模型来处理具有模糊性的现实对象及其关系．这种模型使用的数学思想方法是1965年美国数学家查德提出的模糊数学(包括模糊集合和模糊逻辑)的思想方法，即模糊思想方法．

第四类是突变数学模型．自然界的事物的变化除了连续的、渐变的以外，还有突然发生变化的事物(如地震、突然死亡、水在摄氏零度突然结成冰、材料突然断裂等等)，用前述三类数学模型来处理这类问题并不能真实地反映客观现实．因此，需要建立描述自然界中不连续的突变现象的数学模型．建立这种数学模型所使用的数学思想方法是法国数学家托姆于1968年在《结构稳定性和形态发生学》一书中首先系统地提出的．英国数学家齐曼将这一新的理论定名为"突变论"．1983年，托姆出版了《突变论:思想和应用》一书，进一步阐述了突变理论的基本思想和应用．他经过数学推导，出色地证明了：形形色色的突变现象可用七种基本的突变模型来描述，只要这个突变过程的控制变量不超过四维．这七个基本模型是：尖顶型、折叠型、燕尾型、蝴蝶型、双曲脐点型、椭圆脐点型、抛物脐点型．齐曼等数学家把突变理论广泛应用到力学、物理学、生物学中，还应用到社会科学中，有力地证明了突变数学模型是处理现实世界中突变现象的一个有力工具．

下面我们仅对模糊性数学模型和确定性数学模型各举一例来作进一步的说明．

"老年人"是模糊概念，即是一个模糊集合．查德根据约定的原则给出了"老年人"这个模糊集合的隶属函数式：

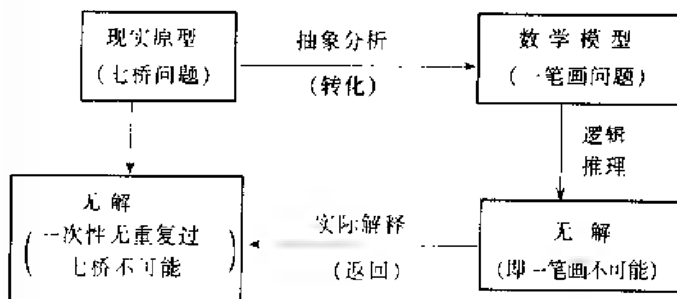
$$
\mu(x)=\begin{cases}0, & x\leqslant 50;\\ \left[1+\left(\dfrac{x-50}{50}\right)^{-2}\right]^{-1}, & x>50.\end{cases}
$$

于是 $\mu(55)=0.5$、$\mu(60)=0.8$、$\mu(70)=0.94$，等等．这表明 50 岁只算“半老”，60，70 岁的人属于“老年人”集合的隶属程度有 0.8，0.94．这样一来，我们就能对本来模糊的语言作了定量的刻画．

本书第一章中谈到的哥尼斯堡七桥问题，则是一个确定性数学模型的例子．

欧拉把这个实际问题，转化为确定的数学模型“一笔画问题”，然后用图论的方法证明了这样的走法根本是不存在的．

现将解决哥尼斯堡七桥问题的过程图示如下：



一般说来，针对一个现实问题，构造一个数学模型并非轻而易举的事情，从上例中我们看到至少应当注意以下几点：

第一，必须对实际问题所属学科领域的基本规律有所了解（当然不一定要求是该领域内的专家，但了解得越多、越深刻越好）．例如，你要构造反映火箭运动规律的数学模型，你就需要熟悉万有引力定律、牛顿第二运动定律等等．如果不熟悉相应领域的基本规律，那就应与相应领域的专家合作来构造数

学模型．如果相应领域里没有现成的理论和方法可供使用，或者虽有但不够用．那就需要自己去观察和探索．总之，对实际问题的背景及内在规律了解得越多、越深刻就越有利构造出有效的数学模型．

第二，要善于抓住实际问题中的本质因素和本质联系，剔除非本质属性，即善于抓住一些起关键性作用的量和相依关系．正如欧拉解决哥尼斯堡七桥问题时所做的那样，桥的长度、宽窄、高低、曲直等都是无关紧要的，小岛的大小、形状也是无关紧要的．

第三，要善于运用数学概念、符号和各种表达式去表现客观现实对象及其相互关系，如果现成的数学工具不够用时，还需要大胆地创建新的数学工具．正如欧拉解决哥尼斯堡七桥问题那样，他仅用点、线这两个最简单的几何元素就把七桥问题的数学模型构造出来，而且为了解决这一问题，欧拉提出新的数学理论和方法，为现代图论这一学科理论的创立奠定了良好的基础．

总之，为了构造好数学模型，必须要有较宽的知识面，要有较强的数学抽象能力和善于运用和创造数学工具．此外，还应将从数学模型上得出的结论返回到实际问题中去检验，由此判断所建立的数学模型是否准确．如果根据实践检验还有一些问题，即与实际不符合，还得修正，有时要经过多次反复，才能使建立的模型更精确地反映客观实际．最后，应该注意到，对于同一个实际问题可能同时建立几个不同的数学模型，其优劣的比较不仅在于对客观现实的描述的精确性上，而且还应该表现在对问题进行分析处理(计算或推理)时是否具有化繁为简、化难为易等特点．

二、强化命题思想

为了获得一个结论，我们往往研究一个更一般的、更强的结论．如果能得出一个更具有普遍意义的结论，那么，原来的问题，作为其特殊情形，自然而然地获得解决．这种用来指导解题的思想称之为强化命题思想．

有的人自然要问：解决一个更一般的问题，证明一个更强的结论，不是更困难、更麻烦吗？通常情况的确是这样，但也不尽如此．这使得强化命题思想有时很有用，对于一些要用数学归纳法证明的命题更是如此．正如前苏联著名数学家辛钦所说的那样："在归纳法证明中，假设命题当 $n-1$ 时成立，再来证它当 n 时也成立．因此，命题越强，在 $n-1$ 的情形下，所给的条件也越多，而对数 n，能证明的东西也越多，但在许多问题中，条件较多显得更为重要．"

例 1 已知 p 是大于 3 的素数，证明 p^2-1 是 24 的倍数．

分析：因为大于 3 的素数都可表成 $6n\pm1$ $(n\in N)$的形式，故改为证明下列加强命题：设 n 为自然数，$p=6n\pm1$，证明 p^2-1 是 24 的倍数．

这一加强命题很容易证明，因而原题也就得到解决．上例中，如果条件不改变，由于素数 p 没有一般表达式供利用，处理起来反而困难得多．

例 2 已知 $0<a<1$，$a_1=1+a$，$a_{n+1}=\dfrac{1}{a_n}+a$ $(n\geqslant1)$，证明对一切自然数 n，有 $a_n>1$．

分析：因为由归纳假设 $a_k>1$，只能推出 $a_{k+1}=\dfrac{1}{a_k}+a<1+a$，而推不出 $a_{k+1}>1$，而要 $a_{k+1}=\dfrac{1}{a_k}+a>1$，必须 $a_k<$

$\frac{1}{1-a}$．故将原命题加强为下列命题：

已知 $0<a<1$，$a_1=1+a$，$a_{n+1}=\frac{1}{a_n}+a\,(n\geqslant1)$，证明：对一切自然数 n，有 $1<a_n<\frac{1}{1-a}$.

加强命题的证明：$n=1$ 时，$1<a_1=1+a=\frac{1-a^2}{1-a}<\frac{1}{1-a}$，设 $1<a_k<\frac{1}{1-a}$，则 $a_{k+1}=\frac{1}{a_k}+a<1+a=\frac{1-a^2}{1-a}<\frac{1}{1-a}$ 且 $a_{k+1}=\frac{1}{a_k}+a>(1-a)+a=1$．即 $1<a_{k+1}<\frac{1}{1-a}$，于是加强命题得证．

上例中如果不将结论加强为 $1<a_n<\frac{1}{1-a}$，则用数学归纳法证题时的归纳假设也较弱，反而不能得出从 $n=k$ 推到 $n=k+1$ 的证明．

例 3 记 a_n 为下述自然数 N 的个数：N 的各位数字之和为 n，且每位数字只能取 1、3、4，求证 a_n 为完全平方数（$n=1$、2，3，…）（1991 年全国数学竞赛高中联赛试题）．

分析：首先依题目条件易得 $a_1=1$、$a_2=1$、$a_3=2$、$a_4=4$ 及 $a_n=a_{n-1}+a_{n-3}+a_{n-4}\,(n\geqslant5)$，但直接由此式来证明题目结论非常困难，我们先将此数列前面若干项计算出来，看有什么一般的规律：

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	…
a_n	1	1	2	4	6	9	15	25	40	64	104	169	273	441	…
规律		1^2	1×2	2^2	2×3	3^2	3×5	5^2	5×8	8^2	8×13	13^2	13×21	21^2	…

通过观察我们得到下列更一般的结论：

加强命题：在原题目的条件下，再设 $f_1=1$、$f_2=2$、f_n

$f_{n-1}+f_{n-2}(n\geqslant 3)$，则有 $a_{2n}=f_n^2$，$a_{2n+1}=f_nf_{n+1}$成立．而这一加强命题不难用数学归纳法证明，从而原命题自然得到证明．

从上述几个例题可以看出，运用加强命题思想解题关键在于构造出一个比原问题更容易解决的加强命题来．如何构造加强命题？并无万能的一般法则，只能针对不同的问题作具体的分析．常用的方法有：1）将所研究的对象放在一个包含它的更大的集合中去考虑(如例 1 中将大于 3 的素数放入形如 $6n\pm1$ 的整数集合中)．2）当进行归纳推理时，需要什么样的条件，而原题又不具备这种条件，则可考虑将它补充到要证的结论中去(如例 2 中为了要证 $a_{k+1}=\frac{1}{a_k}+a<1$ 需要 $a_k<\frac{1}{1-a}$，我们就将它补充到原题要证的结论中去)．3）运用特殊化思想方法，对一系列特殊情形进行观察、分析和比较，归纳出一个更一般的命题(如例 3 所做的那样)．

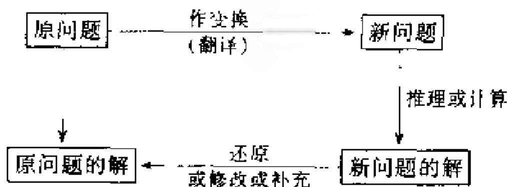
最后，一般化思想在解决数学发展史上遗留的一些著名的难题时，也发挥了巨大的作用．例如古希腊提出的初等几何尺规作图三大难题(三等分任意角、立方倍积、化圆为方)的解决正是这样的：将尺规作图可能性问题一般化，转化成代数方程是否有根式解的问题，而最终运用更一般的数学理论(群论和超越数理论)而加以解决．又如数论中的著名的费尔马大定理(若 $n\geqslant 3$ 为正整数，则方程 $x^n+y^n=z^n$ 无整数解)的解决，也是通过将它一般化，转化成为代数几何中一个更一般性的命题，而用代数几何中的方法加以证明的．可见，采用一般化思想是解决许多数学问题（包括一些数学难题）的一个重要的有效途径．

§5.3 变换思想

人们在解决问题时，对未解决的问题进行一系列变换，将它逐步转化为已知的问题，达到化繁为简、化难为易的目的．这种用来指导解题的思想就是变换思想．

变换如同"翻译"，把同一问题用不同的"语言"，在不同的思维水平上反映出来．如果变换是等价的，那么"翻译"是"真实"的，这时，求出的变换后问题的解就是原问题的解．如解方程、解不等式时应用同解变形定理化简就属于这种情况．如果变换不是等价的，那么"翻译"就会"失真"，这时求出的变换后的问题的解与原问题的解就有一定的差异．在这种情形下，应弄清"失真"的部分及原因是什么，以便对"失真"部分作相应的处理，才能得出原问题的正确解．例如解方程时如果使用了非同解变形去分母、两边平方、两边取对数等等，那么可能产生增根或遗根，就应当将增根舍去而将遗根找回来，才能得到原方程的正确解．

利用变换思想解题的过程可表示为：



利用变换思想解题的关键有二个：一个是变什么？即确定变换的对象；另一个是怎样变？即确定采用什么样的变换．

可以用来作为变换的对象可以是数、代数式、函数、数列、

方程、不等式、几何图形、命题等等，而数式变形、变量替换、几何变换、命题转化等则是解题中常常采用的变换方式.

一、数式变形思想

在解题时，常常要对一些数和式进行变形. 例如，证恒等式时有恒等变形，解方程或不等式时有同解变形等等. 变形的目的是化繁为简、化难为易、化模糊为清晰，从而使问题得到解决. 在初等数学中变形的内容极为丰富，如分解、化简、展开、通分、约分、根式有理化、和差化积与积化和差等等. 在解一个具体数学问题时，应当进行哪些变形和具体如何变形，常常需要一定的技巧(已化成标准型可直接套公式、法则求解的题例外). 我们将不讨论解具体问题的具体变形技巧，仅仅阐述几种用来指导怎样变形的数学思想方法.

(1) 凑配思想

凑是变形中的一种重要思想方法，它指导人们按照题中的目标，对题中出现的数据、式子进行凑合，以期达到某种目的. 例如，凑成能套用公式的形式，凑成能用上题目条件的形式，或直接凑成题目结论所需要的形式等等. 配则包含分配、配对、配方等意思. 配也是变形中一种重要的思想方法，它指导人们在解题时，对题中出现的数据、式子进行合理的搭配或配置，使其易于朝预定的目标进行数式变形. 在实际解题时，凑与配通常是同时应用、交叉进行的.

例 1 设 n 为任意自然数，证明：存在无穷多个正整数 a，使 n^4+a 是合数.

分析：只要选择 a，使 n^4+a 能分成两个整系数因式之积且每个因子之值大于 1 就能完成题目证明. 自然想到取 $a=m^4$(m 为正整数)，于是有 $n^4+a=n^4+m^4=(n^4+2n^2m^2+$

$m^4)-2m^2n^2=(m^2+n^2)^2-2m^2n^2$，如果再继续分解，则系数中出现无理数$\sqrt{2}$，达不到证明目的．为此，应修改我们的"凑法"．将 m 改为$\sqrt{2}m$，即令 $a=4m^4$，则有 $n^4+a=n^4+4m^4=(n^4+4n^2m^2+m^4)-4m^2n^2=(n^2+2m^2)^2-(2mn)^2=(n^2+2mn+2m^2)(n^2-2mn+2m^2)=[(m+n)^2+m^2][(m-n)^2+m^2]$，且当 $m>1$ 时，$(m+n)^2+m^2$ 及 $(m-n)^2+m^2$ 均大于1，从而 n^4+a 为合数，且由 $m>1$ 知有无穷多个 $a=4m^2$ 满足条件．

例 2 设 $I=\{1,2,\cdots,n\}$，对于 I 中的任意非空子集 X，如果 X 中元素从大到小排列为 $a_1>a_2>\cdots>a_k$ $(1\leqslant k\leqslant n)$，那么称 $a_1-a_2+a_3-a_4+\cdots+(-1)^{k-1}a_k$ 为 X 的交替和．试求 X 的一切非空子集的交替和的总和 S．

分析：本题若直接求每一个子集的交替和，再求其总和是很困难的，我们设法将 I 的子集两两配对，而使每一对子集的交替和之和很容易求出，于是所有子集的交替和之总和便可求出．

首先，设空集的交替和为零，并将子集 X 的交替和记为 S_X．其次，将 I 的子集分为二类，第一类中每个子集不含元素 n，第二类中每个子集含元素 n．对于第一类中每个子集 X，我们将它与第二类中子集 $X'=X\cup\{n\}$ 配成一对．显然 $X_1\neq X_2$ 时 $X'_1\neq X'_2$．于是易知 $S_X+S_{X'}=n$，且按这种配对法，I 的所有子集(包括空集有 2^n 个)共配成 2^{n-1}对，故 I 中所有子集的交替和的总和为 $n\cdot 2^{n-1}$，所以 $S=n\cdot 2^{n-1}$（因空集的交替和为零）．

(2) 拆并思想

拆与并指的是分拆与合并，它们是变形中又一个重要的思想方法，在解题中有广泛的应用．例如分解因式时常常要将其

中一项分拆为两项，以便于进行分组分解．又如求某些级数的和 $S_n=a_1+a_2+\cdots+a_n$ 时，常常要将每一项拆成两项的差：$a_k=b_k-b_{k-1}(k=1, 2, \cdots, n)$，于是两两抵消求出 $S_n=b_n-b_0$．然而，如通分、化同次根式、化为同底的对数、化为同角的三角函数等等运算则都是为了达到合并化简的目的．

例 1　设四边形 $ABCD$ 内接于圆，求证 $AB\cdot CD+AD\cdot BC=AC\cdot BD$．

分析：因要证的等式左边有两项，右边只有一项，故应将右边一项拆为两项，再证它们对应相等即可．因此，如图 5-3，应在 AC 上找一点 E 使 $AC\cdot BD=AE\cdot BD+EC\cdot BD$，且满足 $AB\cdot CD=AE\cdot BD$，$AD\cdot BC=EC\cdot BD$．即 $\dfrac{AB}{AE}=\dfrac{BD}{CD}$，$\dfrac{AD}{BD}=\dfrac{CE}{BC}$．也就是只要在 AC 上找 E 使 $\triangle ABE\backsim\triangle DBC$，$\triangle EBC\backsim\triangle ABD$，而这只要作 BE 使 $\angle ABE=\angle DBC$ 即可实现，于是问题得到解决．

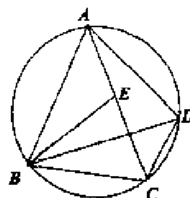


图 5-3

例 2　已知 $a_1, a_2, \cdots, a_n$ 为两两互不相等的正整数，求证 $\sum\limits_{k=1}^{n}\dfrac{a_k}{k^2}\geqslant\sum\limits_{k=1}^{n}\dfrac{1}{k}$．（第 20 届 IMO 试题）

分析：为了便于比较，首先应将两端对应项的分母化为相同，才能合并进行比较．故将 $\sum\limits_{k=1}^{n}\dfrac{1}{k}$ 化为 $\sum\limits_{k=1}^{n}\dfrac{k}{k^2}$，于是，只要证明 $\sum\limits_{k=1}^{n}\dfrac{a_k-k}{k^2}\geqslant 0$ 即可．注意到 $a_1, a_2, \cdots, a_n$ 是两两互不相等的正整数时，必有 $a_1+a_2+\cdots+a_n\geqslant 1+2+\cdots+n$，即有 $\sum\limits_{k=1}^{n}(a_k-k)\geqslant 0$，故要证 $\sum\limits_{k=1}^{n}\dfrac{a_k-k}{k^2}\geqslant 0$，只须将左端各式的分母

逐步化为相同，就能将分式合并，从而达到证明的目的，于是得到下列简捷证法：

$$\sum_{k=1}^{n}\frac{a_k}{k^2}-\sum_{k=1}^{n}\frac{1}{k}=\sum_{k=1}^{n}\frac{a_k-k}{k^2}=\frac{a_1-1}{1^2}+\sum_{k=2}^{n}\frac{a_k-k}{k^2}$$

$$\geqslant\frac{(a_1+a_2)-(1+2)}{2^2}+\sum_{k=3}^{n}\frac{a_k-k}{k^2}\geqslant\cdots\geqslant$$

$$\frac{(a_1+a_2+\cdots+a_n)-(1+2+\cdots+n)}{n^2}\geqslant 0.$$

(3) 增减思想

增包括增加变元个数、增加次数、增加维数、添项等等，减则包括消元、降次、降维、约分等等．运用增的思想指导解题，由于变元的增多，维数、次数的增高，通常会认为将问题变得更为复杂．其实并不完全如此．例如，解方程(组)时，由于引入辅助变量虽然增加了未知数的个数，但却将无理方程(组)化成有理方程(组)，或将高次方程(组)化成低次方程(组)，反而使问题易于求解．用增的思想指导解题的主要作用在于用量的增加的代价去换回质的难度的降低．而运用减的思想指导解题的主要作用是：一方面可以尽量缩小考虑的范围，使条件高度集中，有利于重点突破；另一方面它可以将较复杂的(多元、多因式、高维、高次)数学对象转化为比较简单的(一元、少因式、低维、低次)的情形，从而使问题易于解决．

例 1 已知 $\sin A+\sin B+\sin C=0$，$\cos A+\cos B+\cos C=0$，求 $\cos^2 A+\cos^2 B+\cos^2 C$ 之值．

分析：本题是一个实数范围内的问题，直接在实数范围内，用三角函数的恒等变形来计算很困难，我们不妨将它扩大为复数范围内的问题(升维)来加以解决．

设 $z_1=\cos A+i\sin A$，$z_2=\cos B+i\sin B$，$z_3=\cos C+i\sin C$．由已知条件得 $z_1+z_2+z_3=0$，$\bar{z}_1+\bar{z}_2+\bar{z}_3=0$，$z_1\bar{z}_1$

$=z_2\bar{z}_2=z_3\bar{z}_3=1$，于是 $\cos^2A+\cos^2B+\cos^2C=\frac{1-\cos^2A}{2}+\frac{1-\cos^2B}{2}+\frac{1-\cos^2C}{2}=\frac{3}{2}-(\cos2A+\cos2B+\cos2C)=\frac{3}{2}-\mathrm{Re}(z_1^2+z_2^2+z_3^2)$，而由已知条件可得 $z_1^2+z_2^2+z_3^2=(z_1+z_2+z_3)^2-2(z_1z_2+z_1z_3+z_2z_3)=-2\left(\frac{1}{\bar{z}_1\bar{z}_2}+\frac{1}{\bar{z}_1\bar{z}_3}+\frac{1}{\bar{z}_2\bar{z}_3}\right)=-2\cdot\left(\frac{\bar{z}_1+\bar{z}_2+\bar{z}_3}{\bar{z}_1\bar{z}_2\bar{z}_3}\right)=0$，于是得 $\cos^2A+\cos^2B+\cos^2C=\frac{3}{2}$.

例 2 证明：对一切大于 1 的自然数 n，有 $\left(1+\frac{1}{3}\right)\left(1+\frac{1}{5}\right)\cdots\left(1+\frac{1}{2n-1}\right)>\frac{\sqrt{2n+1}}{2}$.（1985 年全国高考数学试题）

分析：本题困难在于不等式右端很简单，而左端却有 $n-1$ 个因式（多因式），故应设法将一些因式约去（成为少因式）．注意到若将左端记为 A 的话，则有 $A=\frac{4}{3}\cdot\frac{6}{5}\cdot\cdots\cdot\frac{2n}{2n-1}$．为了约去 A 的分子、分母中的因数，我们可令 $B=\left(1+\frac{1}{4}\right)\left(1+\frac{1}{6}\right)\cdots\left(1+\frac{1}{2n}\right)=\frac{5}{4}\cdot\frac{7}{6}\cdot\cdots\cdot\frac{2n+1}{2n}$，于是 $A>B$，$A^2>AB=\frac{2n+1}{3}$，所以 $A>\sqrt{\frac{2n+1}{3}}>\frac{\sqrt{2n+1}}{2}$.

二、变量替换思想

变量替换思想又称为换元思想，其本质之处是用新的变量代换原有的变量或式子，以改变原问题的结构和面貌，达到把复杂和生疏的问题化为简单和熟悉的问题，从而求出原问题的解的目的．

在初等数学中使用的具体代换有很多，例如三角代换、复

数代换、根式代换等等．对于一个具体问题究竟要用什么代换才有效，应针对具体情形作具体分析．这里往往需要丰富的联想和熟练的技巧．但从思想方法的角度来考察，常用代换的主要形式有整体代换、局部代换、分步代换、均值代换等等．

(1) 整体代换

在数式变形中，有时将一个式子用一个字母来替换，使得总体结构简单，关系清晰，便于处理(计算或推理)．我们称这种变量替换为整体代换，其中整体指的是换元的对象．

例 1 设复数 z_1 和 z_2 满足

$$z_1\bar{z}_2+\bar{A}z_1+A\bar{z}_2=0, \quad ①$$

其中 A 是不等于零的复数，证明：

1) $|z_1+A|\,|z_2+A|=|A|^2$;

2) $\dfrac{z_1+A}{z_2+A}=\left|\dfrac{z_1+A}{z_2+A}\right|$.

(1987 年全国高考数学试题)

分析：按常规思路，设 $z_1=a_1+b_1i$，$z_2=a_2+b_2i$，$A=a+bi(a_1, b_1, a_2, b_2, a, b\in R)$，转化为实数集上的问题．由于经过这种代换，出现的字母太多且运算复杂，往往不得不半途而废．但是，如果我们从整体入手作代换 $\alpha=z_1+A$，$\beta=z_2+A$，则已知条件①变为：

$$\alpha\bar{\beta}=|A|^2. \quad ②$$

要证结论化为：

1) $|\alpha|\cdot|\beta|=|A|^2$, 2) $\dfrac{\alpha}{\beta}=\left|\dfrac{\alpha}{\beta}\right|$.

这便是一道十分容易的题了．事实上，由②有 $|\alpha\beta|=|\alpha\bar{\beta}|=|A|^2$，且由 $A\neq0$ 得 $\beta\neq0$，于是 $\dfrac{\alpha}{\beta}=\dfrac{\alpha\bar{\beta}}{\beta\bar{\beta}}=\dfrac{|A|^2}{|\beta|^2}$ 为正实数，所

以 $\frac{\alpha}{\beta}=\left|\frac{\alpha}{\beta}\right|$.

例 2 设$\triangle ABC$的外接圆半径为R，内切圆半径为r．求证$R\geqslant 2r$．

分析：在第三章我们曾用构造图形的方法论证此题．如果熟悉欧拉定理$d^2=R^2-2Rr$（d为外心与内心之间的距离）要证明本题自然无困难，但对于不熟悉这一定理的中学生来说，解决本题并不是轻而易举的．然而，只要设$\triangle ABC$的面积为Δ，半周长为p，边长为a，b，c，我们作如下整体代换：$x=b+c-a$，$y=c+a-b$，$z=a+b-c$，则易算出$\Delta=\sqrt{p(p-a)(p-b)(p-c)}=\frac{1}{2}\sqrt{xyz(x+y+z)}$，$R=\frac{abc}{4\Delta}=\frac{(x+y)(y+z)(z+x)}{8\sqrt{xyz(x+y+z)}}$，$r=\frac{2\Delta}{a+b+c}=\frac{1}{2}\sqrt{\frac{xyz}{x+y+z}}$．于是$R\geqslant 2r\Leftrightarrow\frac{(x+y)(y+z)(z+x)}{8\sqrt{xyz(x+y+z)}}\geqslant\sqrt{\frac{xyz}{x+y+z}}\Leftrightarrow(x+y)(y+z)(z+x)\geqslant 8xyz$．但$(x+y)(y+z)(z+x)\geqslant 2\sqrt{xy}\cdot 2\sqrt{yz}\cdot 2\sqrt{zx}=8xyz$，故$R\geqslant 2r$成立．

从上面例题可以看出，经过整体代换，有时可使结构面貌简化从而使问题变得容易（如例 1），而有时虽然使结构面貌表面上复杂化了，但结构中各元素之间的关系却更加明朗清晰，便于找到解决问题的途径（如例 2）．

（2）局部代换

所谓局部代换指的是仅将问题中一部分式子用新变量代替而作的代换．作局部代换虽然会使变量增多、方程（条件）增多，但可使结构简化、清晰或降低变元的次数，常常给解题带来方便．

例 解方程 $5x^2-x\sqrt{5x^2-1}+x-2=0$.

分析：解无理方程通常采用移项、两边乘方，从而化去根号的方法来解，这将导致解高次方程，求解比较困难．为了化去根号，我们可采用作局部代换的方法，即令 $y=\sqrt{5x^2-1}$，将原方程化为下列方程组：

$$\begin{cases} y^2+1-xy+x-2=0, \\ y=\sqrt{5x^2-1}. \end{cases} \quad 即 \begin{cases} (y-1)(y-x+1)=0, \\ y=\sqrt{5x^2-1}. \end{cases}$$

而此方程组不难化成两个简单的方程组，于是，原方程的解就很容易求出了．

(3) 分步代换

有些较复杂的问题，常常需要连续不断地作代换，才能使问题简化，从而发现其解题途径．

例 已知 $a_1=1$，$a_{n+1}=\frac{1}{16}(1+4a_n+\sqrt{1+24a_n})(n\geqslant 1)$，求 a_n．（1986 年前西德数学竞赛试题）

分析：为了结构的简化应去掉根号，自然令 $b_n=\sqrt{1+24a_n}$，即 $a_n=\frac{1}{24}(b_n{}^2-1)$，代入原递推关系后化简可得 $2b_{n+1}=b_n+3$，再作代换 $b_n=c_n+k$（k 为待定常数），代入得 $2c_{n+1}=c_n+3-k$．为了便于求出 c_n，自然取 $k=3$，即作代换 $b_n=c_n+3$ 后，便得出 $c_{n+1}=\frac{1}{2}c_n$，故 c_n 是公比为 $\frac{1}{2}$ 的等比数列，其通项很容易求出．于是，相应地 b_n 及 a_n 也就容易求出了．

(4) 均值代换

有的问题中同时出现了两个代数式 A 和 B，我们可以作代换 $y=\frac{1}{2}(A+B)$，$A=y+\delta$，$B=y-\delta$，其中 $\delta=\frac{1}{2}(A-$

B)．于是，可以将原问题化为 $y+\delta$，$y-\delta$ 这样对称的形式来处理．如果 δ 较简单甚至为常数，那么常常可使原问题简化．对于含多个代数式的问题，可类似地作平均值代换，有时会给解题带来许多方便．

例 求方程 $(10x+13)^2(5x+8)(x+1)=\frac{1}{2}$ 的实根．

分析：为了使平均值更接近每一个代数式，先把各代数式中 x 的系数变相同，有

$$(10x+13)^2(10x+16)(10x+10)=10\ .$$

作均值代换 $y=\frac{1}{4}(10x+13+10x+13+10x+16+10x+10)=10x+13$，原方程化为 $y^2(y+3)(y-3)=10$，即 $(y^2-10)(y^2+1)=0$，可求出实根 $y=\pm\sqrt{10}$，于是得原方程的实根为 $x=\frac{1}{10}(-13\pm\sqrt{10})$．

三、几何变换思想

几何变换思想在现代几何理论和应用上都发挥着巨大的作用．德国数学家克莱因把几何学看着是研究在变换群下图形的不变量和不变性质的学科，在不同的变换群下，便得到不同的几何学．

什么是几何变换呢？几何变换就是平面到自身的一个一一映射．设在这个映射下，图形 F 对应着图形 F'，若 A 是 F 的任意一点，通过建立的变换对应着图形 F' 中的点 A'，则点 A' 叫做点 A 的象，点 A 叫做 A' 的原象．

若仅改变图形的位置，而不改变图形的形状和大小，这种变换称为合同变换；若不仅改变图形的位置，而且改变大小，但形状不变，这种变换称为相似变换．

1. 合同变换

在平面到自身的一一映射下，如果任意线段的长和它的象的长总相等，那么这种几何变换称为合同变换.

在合同变换下，共线点变成共线点，射线变成射线，平行线变成平行线，角、三角形、多边形和圆分别变为与它们全等的角、三角形、多边形和圆．因此，两点间的距离、角度、线段的长度、面积等都是合同变换下的不变量.

常见的合同变换有反射、平移和旋转三种.

(1) 反射

在平面到自身的一一映射下，若任意一点 A 与它的象 A' 的连接线段，都被定直线 l 所垂直平分，则这种变换叫做关于直线 l 的反射或对称，直线 l 叫做反射轴(或对称轴)，点 A' 叫做点 A 关于轴 l 的对称点.

显然，对称轴上的点在反射变换下是不动点．对称轴及垂直于对称轴的直线在反射变换下是不动直线．在反射变换中，任何一条直线与它的对应直线或相交于对称轴或互相平行，且对应线段的长度总相等，因此反射是一种合同变换.

例 1 (蝴蝶定理)　设 M 是⊙O 中弦 AB 的中点，CD 和 EF 分别是过 M 点的两条弦，连接 DE 和 CF 分别交 AB 于 P 和 Q 两点，则 $PM=MQ$.

分析：如图 5－4，以 OM 为对称轴作 MC 的对称图形 MC'，要证 $PM=MQ$，只要证 $C'P$ 与 CQ 关于 OM 对称，故只需证$\angle 5-\angle 6=\angle 7$，从而这只要证 C'、M、P、E 共圆，但这可从$\angle C'EP+\angle C'MP=\angle C'EP+\angle 4=\angle C'EP+\angle 3=180°$得出.

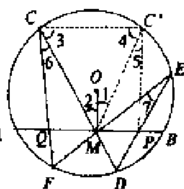


图 5－4

(2) 平移

在平面到自身的一一映射下，若每一个点 A 到它的象点 A' 所连接的有向线段 $\overrightarrow{AA'}$ 等于定向量 $\vec{a}$，则这种变换叫平移，$\vec{a}$ 的方向叫平移方向，$\vec{a}$ 的长度叫平移距离．

平移没有不动点，但有无穷多条不动直线，即平行于平移方向的所有直线．

在平移下，若 A，B 的象点分别是 A'，B'，则 $AA' \mathrel{\underline{\parallel}} BB'$，从而 $AB \mathrel{\underline{\parallel}} A'B'$．这表明平移也是一种合同变换．

例 2　设 P 是▱$ABCD$ 内一点，使得 $\angle PAB = \angle PCB$．证明：$\angle PBA = \angle PDA$．

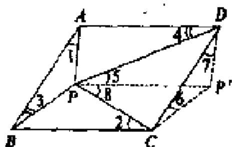


图 5－5

分析：如图 5－5，将 AP 平移到 DP'，则已知 $\triangle APB \cong \triangle DP'C$，于是 $\angle 1 = \angle 7$，$\angle 3 = \angle 6$，要证 $\angle 3 = \angle 4$，只要证 $\angle 6 = \angle 4 = \angle 5$，从而只要证 D，P，C，P' 共圆，但这可由 $\angle 7 = \angle 1 = \angle 2 = \angle 8$ 立即证出．

（3）旋转

在平面到自身的一一映射下，若任意一点 A 和它的象点 A' 到平面上一定点 O 的距离总相等，且 $\angle AOA'$ 等于定角 θ，这种变换叫做关于点 O 的旋转，点 O 叫做旋转中心，θ 叫做旋转角．

在旋转下，除中心外没有其他不动点，一般情形下也没有不动直线（除非旋转角等于 $180°$ 的整数倍）．当旋转角等于 $180°$ 时，每对对应点 A，A' 与中心 O 在一条直线上且线段 AA' 被 O 平分，这种特殊的旋转叫做关于点 O 的中心对称，点 A' 叫做点 A 关于中心 O 的对称点．在中心对称下，过中心的直线是不动直线．设在旋转变换下，A，B 的象点分别是 A'，B'，

则易证$\triangle AOB \cong \triangle A'OB'$，从而$AB = A'B'$，可见旋转也是一种合同变换．

例3 如图5－6，在正方形$A_1A_2A_3A_4$内任取一点P，由A_1向A_2P引垂线l_1，由A_2向A_3P引垂线l_2，由A_3向A_4P引垂线l_3，由A_4向A_1P引垂线l_4．试证：直线l_1，l_2，l_3，l_4交于一点．

分析：本题图形复杂，初看起来似乎无从下手，然而正方形是中心对称图形，绕其中心O顺时针方向旋转90°时，A_1，A_2，A_3，A_4分别变成了A_4，A_1，A_2，A_3，而直线PA_2，PA_3，PA_4，PA_1分别变成了l_1，l_2，l_3，l_4，既然PA_1，PA_2，PA_3，PA_4四线相交于一点P，故直线l_1，l_2，l_3，l_4也相交于一点P'．

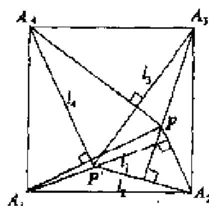


图5－6

2．相似变换

在平面到自身的一一映射下，若任意线段$A'B'$是线段AB的象，且$A'B':AB = k$（k为正常数），则这种变换叫做相似变换，k叫做相似比．在相似变换下，若图形F变为图形F'，则称图形F与图形F'相似，记为$F \backsim F'$．在相似变换下，对应角相等、对应线段长度的比等于相似比，对应图形面积的比等于相似比的平方．显然，合同变换是相似变换的特殊情形，即相似比$k=1$的情形．

位似变换是一种特殊的相似变换，它满足下列条件：1）过对应点的直线AA'通过同一点O；2）$OA':OA = k$（$k \neq 0$为常数）；3）若$k>0$，则A与A'在O同侧，若$k<0$则A与A'在O异侧．

位似变换除了具有相似变换的所有性质外，还有下列性质：在位似变换下，对应直线互相平行。

例 将平面上每个点都以红、蓝两色之一着色，证明：存在这样的两个相似三角形，它们的相似比为1995，并且每一个三角形的三顶点同色。（1995年全国数学竞赛高中联赛试题）

分析：我们用位似变换来作出符合题意的两个相似三角形。因为五边形的5个顶点分别染成两色之一后，必有3个顶点同色，故只要作两个相似五边形使其中一个五边形的5个顶点都同色即可。具体作法如下：以定点 O 为中心，分别以1和1995为半径作两个同心圆 C_1 和 C_2。在 C_1 上任取9点，它们只能染成红、蓝两色之一，故其中必有5点同色，设 C_1 上同色的五点为 A_1，A_2，A_3，A_4，A_5，作射线 OA_1，OA_2，OA_3，OA_4，OA_5 与圆 C_2 为别交于 B_1，B_2，B_3，B_4，B_5 五点，于是，这五点中必有三点同色，设它们是 B_i，B_j，B_k $(1\leqslant i<j<k\leqslant 5)$，于是 $\triangle A_iA_jA_k$ 与 $\triangle B_iB_kB_j$ 相似，相似比为1995且每一个三角形的三顶点同色。

上面介绍了初等几何中常用的几种几何变换，并举例说明了它们的应用。从上述例题我们可以看出，几何变换的主要作用在于通过变换将原来关系不太密切的几何元素（点、线、角等等）和不明显的几何关系（共点、共线、平行、垂直、相等等等）集中在一起，便于发现它们之间的内在联系，从而找出解决问题的途径。

四、命题转化思想

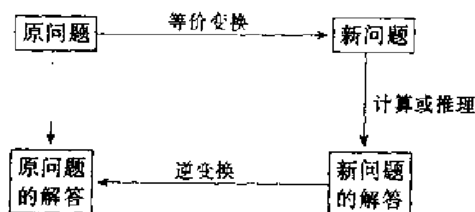
人们在解决问题时，为了寻找解题途径，常常要将一个命题的条件或结论作适当的变化，构造一个新的命题。新的命题与原命题紧密相关，或等价、或较强、或较弱、或相反，通过

对新命题的研究与解决以探求原命题的解决，这种用来指导解题的思想就是命题转化思想．实现命题转化的常用的途径有等价变换、分析与综合、反证及反求等等．

（1）等价变换

如果命题 A 成立可推出命题 B 成立，反过来命题 B 成立可推出命题 A 成立，我们就称命题 A 与命题 B 等价，记为 $A \Leftrightarrow B$ ．

利用等价变换解题的过程可用下列框图表示：



因为一个命题包含条件和结论两部分，所以进行等价变换时，可以仅对条件作等价变换，也可以仅对结论作等价变换，有时则要对条件和结论同时作等价变换，即对整个命题作等价变换．

怎样实现等价转化？常用的思想方法有以下四种：

（i）改变观察问题的角度

有些数学问题，根据题目条件和结论提供的信息，按常规思路去解决问题比较困难，常常需要突破思维定势，换一个角度去思考，改变观察问题的方向，往往能从另一侧面发现问题的本质，从而找出解决问题的办法．

例 1 在 1984 年数学教育国际讨论会上，有一位英国人提出了下列问题：有 4×6 的方格，要在其中 18 个小方格内放置 18 个奶瓶，每格内放一个，但要求每行、每列内放置的奶

瓶数都是偶数．问应如何放？

分析：本题若直接试探将 18 个奶瓶逐一放入，则总会顾此失彼，满足不了每行、每列的奶瓶数都为偶数的要求．如果我们换一个角度，从“不放奶瓶”的角度去思考：在 6 个方格内打“×”(表示不放奶瓶，其余 18 个方格内放奶瓶)，且使每行、每列中打“×”的方格数为偶数，那么问题就容易得多了．如图便是一种放法．

×	×				
×		×			
	×	×			

例 2　集合 A，B 的并集 $A\cup B=\{a_1,a_2,a_3\}$，当 $A\neq B$ 时，(A,B) 与 (B,A) 视为不同的对，则这样的 (A,B) 对共有多少个？(1993 年全国高中数学竞赛联赛试题)

分析：按常规思路(正是竞赛组委会所提供的参考答案中的思路)，应按 A，B 中所含元素的不同情形(A 为空集、一元素集、二元素集、三元素集，再讨论 B 中含元素个数情形)分类进行计算，才能求出正确答案．这样做使解题变为复杂，计算量大，很容易出错．如果我们不从 A，B 包含哪些元素的角度去考虑，而改为考察 a_1，a_2，a_3 这三个元素属于 A，B 的可能情形去讨论，则解答变得非常容易．事实上，由 $A\cup B=\{a_1,a_2,a_3\}$ 知每一个元素 a_i 属于 A，B 的情形只有三种可能：①$a_i\in A$ 且 $a_i\in B$；②$a_i\in A$ 但 $a_i\notin B$，③$a_i\notin A$ 但 $a_i\in B$，而一共只有 3 个元素，故组成的 (A,B) 对共有 $3^3=27$ 种．

(ii) 变更表达问题的方式

有些数学问题，条件或结论的提出比较隐晦、比较抽象或结构比较复杂，我们常常可以换一种方式描述它，使得看起来更为清晰、更为具体或结构更为简单，从而易于触及问题的本质，找到解题的途径．

例 已知$\frac{1}{a}+\frac{1}{b}+\frac{1}{c}=\frac{1}{a+b+c}$，求证$a$，$b$，$c$中必有两个相反数．

分析：本题已知条件用等式表示，而结论却不用式子表示，难以直接推证．我们将结论换成为与其等价的形式：$(a+b)(b+c)(c+a)=0$．这时结论是一个整式等式，而条件却是一个分式形式，结构不一致，为此我们再将条件改为与其等价的形式(去分母)：$(a+b+c)(ab+ac+bc)-abc=0$，由此展开整理便不难得出要证结论．

(iii) 构造一个模型

在论述“一般化思想”的那一节中，我们曾经指出将所研究的问题模型化，是解决实际问题的一种重要思想方法．模型化的实质是结构上进行等价变换，它也是解决数学问题常用的思想方法．通过构造模型常常使问题变得直观，关系更加明朗，从而使问题易于获得解决．

例 证明$\sum\limits_{k=1}^{n}C_n^k\cdot 2^kC_{n-k}^{\left[\frac{n-k}{2}\right]}=C_{2n+1}^n$，这里$\left[\frac{n-k}{2}\right]$表示不超过$\frac{n-k}{2}$的最大整数．(1994年第9届中国数学奥林匹克试题)

分析：本题通常用母函数方法或递推方法去证明，需要一定的变形技巧和较复杂的计算．下面我们用构造组合模型的方法来证明，则显得较简单．

建立模型如下：从$2n+1$个人中选出n个人，共有多少种不同的选法？

一方面从$2n+1$个人中选出n个人有C_{2n+1}^n种方法；另一方面将$2n+1$个人分成n个两人组及一个一人组：

$$\{a_1, a_2\}, \{a_3, a_4\}, \cdots, \{a_{2n-1}, a_{2n}\}, \{a_{2n+1}\} \quad ①$$

对每一个$k(0\leqslant k\leqslant n)$，考虑恰有$k$人来自$k$个2人组，每组

选出 1 人的方法数．因为从①中前 n 组选出 k 组，每组选 1 人的方法有 $C_n^k \cdot 2^k$ 种，其余 $n-k$ 人选自剩下的组．若 $n-k$ 为奇数，则 a_{2n+1} 必选出，其余 $n-k-1$ 人选自剩余的 $\frac{n-k-1}{2}$ 组，有 $C_{n-k}^{\frac{n-k-1}{2}} - C_{n-k}^{[\frac{n-k}{2}]}$ 种方法；若 $n-k$ 为偶数，则 a_{2n+1} 必不选出，余下 $n-k$ 人选自剩余的 $\frac{n-k}{2}$ 组，有 $C_{n-k}^{\frac{n-k}{2}} - C_{n-k}^{[\frac{n-k}{2}]}$ 种方法．可见，不论怎样情况，恰有 k 人选自 k 个 2 人组，每组选 1 人的方法数为 $C_n^k \cdot 2^k C_{n-k}^{[\frac{n-k}{2}]}$，而 k 可以为 0，1，2，…，n，从而共有 $\sum\limits_{k=0}^{n} C_n^k \cdot 2^k C_{n-k}^{[\frac{n-k}{2}]}$ 种选法，所以 $\sum\limits_{k=0}^{n} C_n^k \cdot 2^k C_{n-k}^{[\frac{n-k}{2}]} = C_{2n+1}^n$.

(iv) 选择有效的记号

当有些问题表达非常复杂、叙述非常啰嗦时，常常使人摸不着要领，抓不住问题的本质，感到无从入手．这时如果恰当地引入有效的记号，则往往能将问题中的内在联系清楚而简洁地表示出来．

例 英国数学家哈密尔顿曾提出这样一个问题：在一个图 G（例如图 5－7 那样的图），能否找到一条道路，从一个顶点出发经过所有的顶点，再回到原来出发的顶点，而不走重复的线路．如果存在这样的闭路，就称为具有一个哈密尔顿圈．现在要问：图 5－7 是否有一个哈密尔顿圈？

如果直接对图 5－7 进行验证，将不胜其烦，但如果像图 5－8 那样给每个顶点标上 4 或 3（即该顶点所连接的线的条数），那么从任一标有 3 的顶点只能走到标有 4 的顶点，同样地，从标有 4 的顶点只能走到标有 3 的顶点．如果存在哈密尔顿圈，标 3 的顶点应该和标 4 的顶点一样多．但图 5－8 中有 6 个标 4 的顶点，8 个标 3 的顶点，所以不存在哈密尔顿圈．

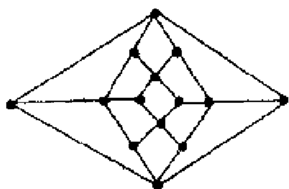


图5-7

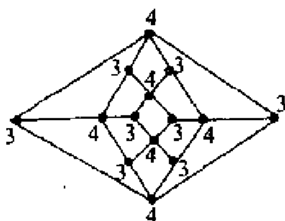


图5-8

我们在上面介绍了四种实现命题等价变换的基本途径，但实现等价变换的手段远远不只这些．读者在研究和解决数学问题时，应善于运用等价变换的思想指导自己去变更所研究的问题．至于采用什么样的等价变换，除了上述的一些方法及本书其他章节介绍的一些方法（如变量替换、赋值、染色、同构等等）外，还有待读者更多地去探索和总结．

其次，进行等价变换，一定要目的明确，如果进行变换后得到的命题比原来更复杂，更难于解决，那么所作的变换也就失去了意义．

最后，等价变换仅是众多变换中的一种．有些问题用等价变换达不到简化问题的目的，而且非等价变换反而能化难为易，化复杂为简单，这时反而要用非等价变换．

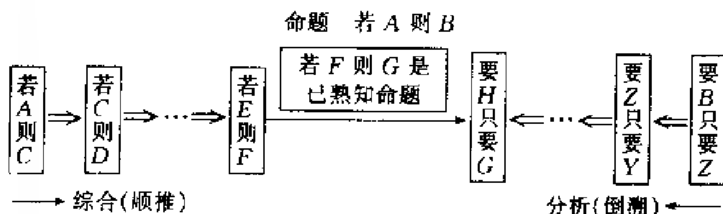
(2) 分析与综合

所谓分析法，就是执果溯因的方法．它是从结论出发，追溯使结论成立需要什么充分条件，再探索为了得到这一条件又需要什么充分条件，最后归结至已知条件或已有的结论（定义、公理或定理）．其一般证法是：要证命题“若 A 则 B”，可证“要 B 只要 Z”，“要 Z 只要 Y”，…，“要 D 只要 C”，“要 C 只要 A”，而“今有 A，故有 B”．因此，分析法是在不断地转化命题，每转化一次就使要证结论向前转化了一次，而结论每

向前转化一次就越接近已知条件或已有结论，最终使问题得到解决.

所谓综合法，就是由因导果的方法．它是从已知条件出发，结合已有的结论进行顺推，逐步推证前提成立的必要条件，最后证得结论成立．其一般证法是：设要证命题"若 A 则 B"，可证"若 A 则 C"，"若 C 则 D"，…，"若 Y 则 Z"，"若 Z 则 B"，于是"若 A 则 B"成立．可见，综合法也是在不断地转化命题，每转化一次就使前提向后转化了一次，使已证出的条件越来越接近要证结论，从而获得原问题的证明.

通常解题时，并不单纯只使用分析法或只使用综合法，而是两种方法同时并用．一方面从前提出发顺推，向结论逼近；另一方面从结论出发倒溯，向前提靠拢，两者汇合到一起成为一个熟知的命题为止．具体过程如下：



还必须注意，用综合法顺推时，不必每次导出一个必要条件，用分析法倒溯时也不必每次倒推出一个充分条件，只要有联系就行．找到思路后，再判断思路是否可行，对于推不出的地方，还应补充证明．这样做才不致于过于拘泥而使思路不开阔，又不致于由于推理不严密造成漏洞百出.

由于分析法、综合法进行的命题转化往往是不等价的，因此，这种转化带有较大的灵活性，不同的转化常常导致不同的

解法.

例 已知△ABC 三内角 A, B, C 成等差数列，求证三边满足$\frac{1}{a+b}+\frac{1}{b+c}=\frac{3}{a+b+c}$.

分析:(i) 由已知条件顺推:

A、B、C 成等差数列$\to 2B=A+C\to B=60^\circ$.

(ii) 由结论倒溯:

要证 $\frac{1}{a+b}+\frac{1}{b+c}=\frac{3}{a+b+c}\to$ 只要证 $\frac{a+b+c}{a+b}+\frac{a+b+c}{b+c}=3\to$ 只要证$\frac{c}{a+b}+\frac{a}{b+c}=1\to$ 只要证 $c(b+c)+a(a+b)-(a+b)(b+c)$即 $a^2+c^2-ac-b^2$.

(iii) 至此，只要利用余弦定理: $b^2=a^2+c^2-2ac\cos B$ 便可完成证明.

(3) 反证

反证法是数学中一种非常重要的思想方法，法国数学家阿达玛曾经说过:"若肯定定理的假设而否定其结论，就会导致矛盾."这是对反证法实质的精辟概括. 具体说来，若肯定命题的假设而否定其结论，经过正确的逻辑推理，导出了矛盾，就证明了"否定其结论"是错误的，从而证明了结论的正确性.

这种证明方法如果用逻辑符号表示，可叙述为: 证明命题"$A\to B$"(若 A 则B)时，如果假设 $A\wedge\overline{B}$($\overline{B}$ 表示结论的反面，$A\wedge\overline{B}$ 表示同时有A 和$\overline{B}$)为真，在 A 和B 同时成立的条件下，如果逻辑地推出一个"$C\wedge\overline{C}$"来，而根据矛盾律 C 与C 不能同真，断定 $A\wedge\overline{B}$ 为假，再根据排中律，两个互相矛盾的判断不能同假，由此肯定命题"$A\to B$"为真.

反证法的逻辑结构可简明地表示成下列等价命题:

$$\text{“}A\to B\text{”}\Leftrightarrow\text{“}A\wedge B\to C\wedge\overline{C}\text{”} \qquad ①$$

现用真值表给出这个等价命题的证明：

A	B	$A\to B$	$\bar{B}$	$A\wedge\bar{B}$	$C\wedge\bar{C}$	$A\wedge\bar{B}\to C\wedge\bar{C}$
0	0	1	1	0	0	1
0	1	1	0	0	0	1
1	0	0	1	1	0	0
1	1	1	0	0	0	1

由上表可知，对命题 A、B 的一切取值组合“$A\to B$”与“$A\wedge\bar{B}\to C\wedge\bar{C}$”均同真同假，这就证明了“$A\to B$”与“$A\wedge\bar{B}\to C\wedge\bar{C}$”等价．

于是，要证明命题“$A\to B$”，可代之证明它的等价命题“$A\wedge B\to C\wedge C$”，这就是反证法．

读者不难用真值表类似地证明命题“$A\to B$”与其逆否命题“$\bar{B}\to\bar{A}$”等价．因此，要证明命题“$A\to B$”，也可用证明它的另一等价命题“$\bar{B}\to\bar{A}$”来代替，这也是反证法的一种形式．实际上证明了“$\bar{B}\to\bar{A}$”后，由“$A\wedge\bar{B}$”便可推出矛盾“$A\wedge\bar{A}$”来．即从结论的否定 $\bar{B}$ 出发导至与已知条件 A 矛盾，从而断言结论 B 成立．但一般说来，用反证法证题导出的矛盾，除了是和已知条件矛盾外，还可以与已有的定理、公理、定义矛盾，或推出两个互相矛盾的结果．可见，将反证法的逻辑结构写成①的形式比写成“$A\to B$”$\Leftrightarrow$“$\bar{B}\to\bar{A}$”这样的形式，应用起来更为方便．

用反证法证题的步骤是：

(i) 反设：假定结论的反面成立；

(ii) 归谬：由反设的假定出发，结合已知条件推出矛盾．这个矛盾可以是推出的结果与已知条件、公理、定理或定义中任何一个的矛盾，也可以是推出的结果与反设或与推出的另一

结果的矛盾；

(iii) 结论：假定结论的反面成立是不正确的，从而肯定结论成立。

在数学中，反证法的应用非常广泛，它被称赞为“数学家的最精良的武器之一”，有关“存在性”、“惟一性”、“否定性”、“无限性”的命题，常常采用反证法，对于一些用直接证法难于奏效的问题，也常用反证法。

例 平面内 n 个点之间共连有 m 条线段，且每条线段只与两点相连。证明：当$m>\frac{1}{2}(n-1)(n-2)$时，任意两点间都有折线相连通。

分析：本题要直接证明任意两点间都有折线相连通是较困难的，故考虑用反证法。假设 n 个点中存在两点A 和B，它们没有折线相连通。并设与 A 有折线相连通的点共有k 个($1\leqslant k\leqslant n-2$)，它们是 A_1、A_2，…，A_k。于是 $k+1$ 个点 A，A_1，…，A_k 之间至多连有C_{k+1}^2条线段，其余 $n-k-1$ 个点之间至多连有 C_{n-k-1}^2条线段，并且其余 $n-k-1$ 个点中任何点与 A，A_1，…，A_k 中任何点没有线段连接(否则，与 A 连有折线的点将多于k 个)，于是 $m\leqslant C_k^2+C_{n-k-1}^2$，而不难证出 $C_k^2+C_{n-k-1}^2\leqslant\frac{1}{2}(n-1)(n-2)$，这与已知条件$m>\frac{1}{2}(n-1)(n-2)$矛盾，故原命题成立。

(4)反求

当我们求一个量 x 的值很困难时，可反过来求与 x 相关的另一个量y，再由 x 与y 的关系式$y=f(x)$，便可求出 x 来。在证明中也有类似的情况，要证 A 有性质B 有困难，可先证 A 没有性质C，而没有性质 C 很容易推出必有B，也就证明

了 A 必有 B .

例 1　求 $f(\theta,\varphi)=A\cos^2\theta+B\cos^2(\varphi+\theta)+C\cos\theta\cos(\varphi+\theta)$ 的值与 θ 无关时，A，B，C 应满足什么充要条件?

分析：按常规思路，应将式中 θ 与 φ 分离开，再令含 θ 的式子等于零，便可求出 A、B，C 应满足的充要条件．但这样做不仅变形技巧繁琐，而且计算量大．不如反过来讨论，既然与 θ 的值无关，可令 θ 取一些特殊值．例如：

$$令\ \theta=0, f(0,\varphi)=A+B\cos^2\varphi+C\cos\varphi\ , \qquad ①$$

$$令\ \theta=\frac{\pi}{2}, 得\ f\left(\frac{\pi}{2},\varphi\right)=B\sin^2\varphi\ , \qquad ②$$

$$令\ \theta=-\varphi，得\ f(-\varphi,\varphi)=A\cos^2\varphi+B+C\cos\varphi\ . \qquad ③$$

因 $f(\theta,\varphi)$ 与 θ 无关，于是 $f(0,\varphi)=f(\frac{\pi}{2},\varphi)=f(-\varphi,\varphi)$．从①，②，③中消去 f 可得 $A=B=-\frac{C}{2}\sec\varphi$，即得必要性，充分性则可直接代入去证明(证明略).

例 2　解方程 $x^3+2(\sqrt{2}+1)x^2+(3+2\sqrt{2})x+\sqrt{2}=0$.

分析：这是关于 x 的三次方程，直接解太繁了！令 $a=\sqrt{2}+1$，原方程化为

$$x^3+2ax+a^2x+a-1=0,$$

即

$$xa^2+(2x+1)a+(x^3-1)=0\ .$$

当成 a 的二次方程易求出 $a=-x+1$ 或 $a=-\frac{x^2+x+1}{x}$，于是不难求出 $x=-\sqrt{2}$，或 $x=-\frac{1}{2}(2+\sqrt{2})\pm\frac{1}{2}\sqrt{2+4\sqrt{2}}$.

第六章　转化思想(下)

在这一章中,我们将继续讨论两种重要的转化思想,即映射思想与数形结合的思想,它们可以归结为第二类的转化思想.映射与数形结合不仅是重要的数学思想,而且还直接作为一种方法、技巧,在数学研究和解题中发生作用.

§6.1　映射思想

比"映射"更基础的概念是"对应",在一般的数学文献中是作为一个不加定义的术语使用的,是一个最基本最重要的数学概念.

几何中的各种变换,代数中的各种运算,分析中的各种函数,都是对应的例子.

组合数学中的许多证明和计数问题,都是利用对应的思想方法来解决的.

在现代数学中,许多重要的概念,如同态、同构、同伦、同胚等等,无一不是某种对应思想的反映.被誉为20世纪最重要数学成果的法尔廷斯(Faltings)定理,就是因证明了一系列的对应都是同构的,从而解决了长期悬而未决的莫德尔(Mordel)猜想.

由此可见,对应思想在数学中的重要.

对应思想也是人类最早期、最容易掌握的一种思维方式.儿童还在牙牙学语的时候,就会使用勾手指头的办法来计数苹果的个数,他们将桌上的苹果与自己的手指头一一对应.中非

森林中的某一原始部落的居民，在19世纪还基本上没有数的概念，却能用大量的象牙与海外商人进行换货贸易：每一包烟草上放置一支象牙，这样就得到两个数量上相等的集合，然后一方收取烟草，另一方收取象牙．如果要以分别计数的方法进行交易，当地人是不能接受的．

我国很早就使用对应思想作为研究的工具了．在我国古老的儒家经典《周易》中，将人世间的万象万物与“八卦”相对应，企图运用“易卦”中的阴阳交错、交位高低等“卦象”来探测宇宙人生的奥秘．古代星相学家将天上的星宿与地上的人事相对应，通过对天象的分析来推测人世的动态．我国古代的五行学说则将万事万物分别与金、木、水、火、土等“五行”相对应，力图利用五行“相生相克”的规律来演绎事物之间的联系和变化，等等．

由此可见，对应不仅是一种重要的数学思想，也是人类最早期掌握、最普遍使用的一种思想方法．

一、映　射

1．序偶·关系·映射

1921年，波兰数学家库拉托夫斯基研究了对应的结构关系，把它抽象出来，给出了“序偶”的概念．

设A与B是两个集合，A的元素用a表示，B的元素用b表示，在A中取元素a，B中取元素b，将a与b排成顺序后，称为集合A与集合B之间的一个序偶．

例如，设集合$A=\{1, 2, 3\}$，$B=\{$奇，偶$\}$，则(1，奇)，(2，偶)，(3，偶)等都是A与B的序偶．但(1，2)，(偶，2)等都不是A与B的序偶，因为(1，2)中的第二个元素不是B的元素，(偶，2)中B的元素在A的元素之前，只是B与A的一个序偶，而不是A与B的一个序偶．

把 A 与 B 的全部序偶当作元素作成一个新集合，则称为 A 与 B 的一个笛卡儿积，记作 $A\times B$.

例如，集合 $A=\{1, 2, 3\}$ 与集合 $B=\{$奇，偶$\}$ 的全部序偶是：

(1,奇)，(1,偶)，(2,奇)，(2,偶)，(3,奇)，(3,偶)．

所以 A 与 B 的笛卡儿积是：

$A\times B=\{$(1，奇)，(1，偶)，(2，奇)，(2，偶)，(3，奇)，(3，偶)$\}$．

集合 A 与 B 的笛卡儿积 $A\times B$ 的一个子集 R（或者说 A 与 B 的一部分序偶所成之集）叫做 A 与 B 的一个二元关系．

例如，上面谈到的 $A\times B$ 的子集

$R=\{$(1，奇)，(2，偶)，(3，奇)$\}$

是 A 与 B 的一个二元关系．这个关系反映了对集合 A 中的数按奇偶性分类的作用．

当 $A\times B$ 的元素 $(a, b)\in R$ 时，我们说 a 与 b 具有关系 R；当 $(a, b)\notin R$ 时，则说 a 与 b 没有关系 R .

二元关系可以看作许多实际问题的数学模型，也可以看作许多数学概念的抽象结构．

有了序偶和关系的概念，就可以给出映射的概念了．

定义　设 f 是 A 与 B 的一个二元关系，如果对于每一个 $x\in A$，都有惟一的 $y\in B$，使序偶 $(x, y)\in f$，则称 f 是 A 到 B 的一个映射．记作：

$$f: A\longmapsto B$$

在初等数学中，一般都避开序偶和关系的概念，直接使用"对应"来定义映射：

定义　设 A 与 B 是两个集合，对每一个 $x\in A$，在 B 中都可以通过某种固定的法则找到惟一确定的元素 y 与之对应，则称 f 是 A 到 B 的一个映射．记作

$$f:A \longmapsto B.$$

其中的 A 称为原象集合，A 中每一个元素 x 称为映射的原象，它对应的元素 y 称为 x 在映射 f 下的象．记作

$$f: x \longmapsto y,$$

或

$$y=f(x).$$

由映射的定义可知，数学中的各种函数是一种特殊的映射，平面上的点 A 对应它的坐标 (x,y) 是一个映射，将所有三角形中每一个与它的面积相对应，也是一个映射．

映射的具体内容和方式是多种多样的，但从映射的结构分析，又有满射和单射之分．

设 f 是 A 到 B 的一个映射．

如果对于集合 A 中任意两个不同的元素 x_1 和 x_2，当 $x_1 \neq x_2$ 时，都有

$$f(x_1) \neq f(x_2),$$

即 A 中不同的元素在 B 中的像也不同，则称 f 是 A 到 B 的单射．

如果对于集合 B 中任一元素 y，都有 A 中的元素 x（至少一个），使得

$$f(x)=y,$$

即集合 B 中每个元素都是集合 A 中某个或某些元素的象，则称 f 为 A 到 B 的满射．

如果映射 f 既是满射，又是单射，则称 f 为 A 与 B 之间的一个一一映射．

如果 f 是 A 到 B 的一个一一对应，且

$$f: x \longmapsto y,$$

作新的对应：

$g: y \longmapsto x,$

则 g 是集合 B 到集合 A 的一个映射，称为 f 的逆映射，通常记作 f^{-1}. 显然 f^{-1}也是一一对应.

人们特别重视一一对应的映射. 俗话说的“一个钉子一个眼，一个萝卜一个坑”，就是对一一对应的一个通俗而形象的概括.

2. 映射的作用

映射之所以成为一种重要的数学思想，可以从一些简单的事例中体现出来.

曾经在比萨斜塔上做过著名的重力实验的意大利物理学家伽利略就曾经利用映射发现了一个有趣的现象，被称为“伽利略的悖论”.

平方数只是自然数的一部分，“整体大于部分”应该是颠扑不破的真理. 但是伽利略发现自然数与完全平方数之间可以建立一一对应的关系：

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & \cdots & n & \cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \cdots & \updownarrow & \cdots \\ 1 & 4 & 9 & 16 & 25 & \cdots & n^2 & \cdots \end{array}$$

有一个自然数，就有一个相应的平方数.反之，有一个完全平方数，也就有一个相应的自然数. 既然在两个集合 A 与 B 中，A 有的每一个数 B 都有相应的数，B 有的每一个数 A 也都有相应的数，并且不同的数相对应的数也不同，两者不是“一样多”吗？全体和部分“一样多”，简直是不可思议！伽利略的怪论惊呆了当时的许多人.

伽利略本人未能解开这一个谜，直到他逝世200多年以后，德国数学家康托建立了无穷集合的理论，才彻底解决这个问

题．“全体大于部分”的公理只在有穷集中才能成立，在无穷集中并不成立．从此，人们对无穷集的本质有了新的理解：

“一个可以和自己的某个真子集建立一一对应的集合称为无穷集．”

从此，使人们对无穷集理解跳出了那种“无穷就是取之不尽”之类的同义语的圈子．由此可见，映射对发现数学的内在矛盾，认识数学问题的本质有着非常重要的作用．

在组合数学中，解决有关计数的问题时，一个最基本的方法就是利用映射来转化被计数的对象．

例如有这样一道题，假设有100名运动员参加乒乓球单打比赛，比赛方法采用淘汰制，把运动员分成两两一组进行比赛，败者被淘汰，胜者进入第二轮．再把进入第二轮的运动员两两分组进行比赛，败者被淘汰，胜者进入第三轮．如此继续下去，如果遇到某一轮的运动员为单数，则令其中一名运动员轮空直接进入下一轮．最后决出冠军，问一共要进行多少场比赛？

这是一道很简单的数学题，小学生都会解答，但他可能这样来计算：

第一轮 100 人分成 50 组，进行 50 场比赛；

第二轮 50 人分成 25 组，进行 25 场比赛；

第三轮 25 人一人轮空分成 12 组，比赛 12 场；

第四轮 13 人一人轮空分成 6 组，比赛 6 场；

第五轮 7 人一人轮空分成 3 组，比赛 3 场；

第六轮 4 人分成 2 组，比赛 2 场；

第七轮 2 人进行 1 场比赛决出冠军．

因此，总的比赛场数是：

$$50+25+12+6+3+2+1=99\text{．}$$

这是一种不太高明的算法，当参赛的运动员人数较多时，计算是比较麻烦的。特别是如果参赛人数为一般的 n 时，则无法按上面的方法进行计算。

数学家对这个问题则会利用映射的思想。

每一场比赛淘汰惟一的一名运动员；反之，每一名被淘汰的运动员在惟一的一场比赛中被淘汰。因此，被淘汰的运动员的集合与比赛场次的集合之间可以建立一一对应的关系。假设最初有 n 名运动员参加比赛，不管 n 是多少，最后只有1名冠军，其余的 $n-1$ 名运动员均被淘汰，所以恰好要进行 $n-1$ 场比赛。

这后一种解法无须作任何计算，而且具有更深刻、更本质的意义。

由此可见，映射思想在优化、简化解题方法方面有着重要的作用。下面我们选择几种重要的映射思想加以论述。

二、配对思想

配对既是一种重要的数学思想，又是完成映射的一种重要方法。配对思想的特点是，建立集合 A 与另一集合 B 的元素之间的特殊映射关系，使得原像和它的像之间具有某些特殊的性质，如同一、凑整、抵消、共轭、对称、互补、互余等等，使它们或相辅相生，或相反相成，从而为解决问题提供某些有利的条件。

1. 倍数映射

我国历史上有许多以弱胜强的著名战例，齐魏马陵之战就是其中之一。公元前344年，魏国大举进攻韩国，齐王派田忌和孙膑率兵救援韩国。孙膑并不直接出兵韩国，却采取了“围魏救赵”的策略，驱军直逼魏国的都城。魏将庞涓不得不回师

迎战齐军．魏军一路追击齐军，庞涓看见齐军宿营地的灶一天天在减少．根据灶的数目可以推算出齐军的数目，断定齐军已经严重减员．便轻骑冒进，猛打穷追．殊不知这原来是孙膑的"增兵减灶"之计，结果庞涓中了齐军的埋伏，全军覆灭．

如果不是孙膑制造假象设下陷阱，庞涓的思想是正确的．他正是运用了一种倍数映射的思想，因为灶与士兵之间可以建立一种倍数映射关系．例如若一口灶供10人造饭，齐军有 3 万灶，便可推知齐军有 30 万人．

倍数映射是数学中解决计数问题的常用方法，也是一种常见的配对思想．

假定我们要计数的集合 A 很复杂，不容易直接计数，我们便设法找到另一个易于计数的集合 B，使 B 中每 n 个元素与 A 中的一个元素对应，换言之，使 A 中的元素与 B 中的 n 元素组建立一一对应关系，就称为 A 与 B 之间的一个倍数映射．

特别地，当 $n=1$，就是通常的一一映射．

设集合 A 的元素个数为 $N(A)$，集合 B 的元素个数为 $N(B)$，则有

$$N(B)=nN(A).$$

利用这一公式可以解决许多计数问题．

例 1 求用 1，2，3，4 所组成的一切三位数之和：

1）假定数字不可重复使用；

2）假定数字可以重复使用．

我们用 B 表示由1，2，3，4四个数字所组成的三位数所组成的集合．因为1，2，3，4可以这样组成配对，使每对之和均为 5，所以集合 B 中任一元素 b 都可以在 B 中找到另一元素 b'，使 $b+b'=555$．比如取 $b=413$，则可取 $b'=142$．现在

我们用 A 表示由集合 B 中 b 与 b' 组成的数对的集合．即

$$A=\{(b,b')\mid b,b'\in B,\quad b+b'=555\}.$$

于是我们在 A 与 B 之间建立了一个 2 倍映射．下一步我们求出 A 中元素个数 $N(A)$，再乘以555就可以了．根据公式

$$N(B)=2N(A).$$

在第1)问中，B 的元素不能使用重复数字，故 $N(B)=P_4^3=24$，从而 $N(A)=12$，所以第 1) 问的答案是：

$$555\times 12=6660.$$

在第2)问中，B 的元素可以使用重复数字，故 $N(B)=4^3=64$，$N(A)=32$，第2)问的答案是：

$$555\times 32=17760.$$

这个问题实际上应用了两次配对，第一次 b 与 b' 配对的一一映射，第二次 (b,b') 与 a 配对的倍数映射．

例 2 如图6－1，一个边长为 3 的正方体，将它横切两刀，纵切两刀，再竖切两刀，可以切成 27 个边长为 1 的单位正方体．如果允许把每次切出的小块任意叠起来切，能否不用 6 刀(例如 5 刀或更少)就能得出 27 个单位立方体呢?

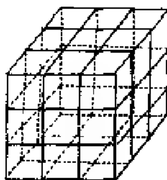


图 6－1

问题毋须试验．关键在于位于中心的那个小立方体，不管你怎么切，每切一刀只能使它的一个面"曝光"．因此，切的刀数与中心那个小立方体的面可以一一配对，即建立一一对应关系．小正方体有6个面，所以，一定要切 6 刀才能切出 27 个单位立方体．

例 3 设有一个凸 n 边形，在它的内部没有任何三条对角线交于一点，求它的所有对角线在它内部交点的总数．

如图 6－2，多边形的两条对角线 AB、CD 决定一个交点

H. 反过来，一个交点 H 对应两条对角线 AB 和 CD，或者说对应一个由 n 边形的 4 个顶点所成的组 (A,B,C,D). 因此，若记

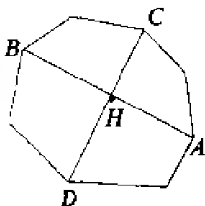


图 6－2

$T=\{H|n$ 边形的对角线在内部的交点$\}$,

$S=\{(A,B,C,D)|n$ 边形的 4 顶点组$\}$,

则 T 与 S 之间可以建立一一对应关系，所以有

$$N(T)=N(S).$$

$N(S)$ 是很容易计算的，它是 n 个元素中取 4 个元素的组合数，即 C_n^4.

例2、例 3 都是一一映射的例子. 一一映射不仅是解决计数问题的重要思想方法，也是证明某些组合不等式的重要思想方法. 这一思想方法的特点是把要证的不等式转化为比较两个集合 A 与 B 的元素的多少. 欲证 $N(A)<N(B)$，可建立 A 与 B 的一个子集之间的一一对应或证明有一个 A 到 B 的一个子集之间的内射而不是满射.

例 4 用 $2n(n>1)$ 条直线把平面分成若干个区域，假设这些直线任意两条不平行，任意三条不共点. 证明在分成的这些区域中，角状区域不能多于 $2n-1$ 个.

利用递推方法可以求得 n 条直线划分平面的区域数 a_n 满足递推公式

$$a_n=a_{n-1}+n,$$

但这一公式中并未能提供角状区域有多少的信息，因此利用递推的方法并不方便，利用映射则较为方便.

因为 $2n$ 条直线只能有有限个交点，故可作一充分大的圆使所有交点都包含在圆内. $2n$ 条直线将这个圆分成 $4n$ 段圆

弧．显然，每个角状区域内部含有一段圆弧，但两段相邻的圆弧不能都属于角状区域，否则将出现三条线共点的情况．因此我们可以建立角状区域的集合到$2n$段圆弧的集合之间的一个单射，因而由$2n$条直线划分平面所得的角状区域数b_n满足不等式：

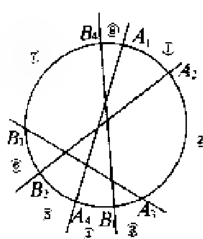


图6－3

$$b_n \leqslant 2n\ .$$

为了证明这个映射不是满射，可将$4n$条圆弧，从某一角状区域所含的一段圆弧开始，按顺时针编号为1，2，…，$4n$．则编号为奇数的圆弧才能属于角状区域，但很明显，编号为$2n+1$的那段圆弧不能也属于角状区域．这就证明了这个映射不是单射．所以

$$b_n < 2n-1\ .$$

例5　设n是正整数，我们说集合$\{1,2,\cdots,2n\}$的一个排列$(x_1,x_2,\cdots,x_{2n})$具有性质P，如果在$\{1,2,\cdots,2n-1\}$中至少有一个i，使得$|x_i-x_{i+1}|=n$．求证：对于任何n，具有性质P的排列比不具有性质P的排列多．（第30届IMO试题）

要证明一个如此复杂的不等式，用普通方法显然有一定的难度．无怪当年各国领队在讨论是否采用这道试题时，有的人认为太难，也有的人认为不难，争执不下．竞赛委员会主席最后幽默地说："难还是不难，让参加竞赛的学生们去判断吧！"结果有的学生巧妙地运用了映射的方法，得出了十分简捷的证明：

记$A=\{(x_1,x_2,\cdots,x_{2n})\mid(x_1,x_2,\cdots,x_{2n})$具有性质$P\}$，

$B=\{(x_1, x_2, \cdots, x_{2n}) \mid (x_1, x_2, \cdots, x_{2n})$ 不具有性质 $P\}$，

$C=\{(x_1, x_2, \cdots, x_{2n})$ 恰有某一个 i 使 $|x_i-x_{i+1}|=n$，但 $i\neq 1\}$。

显然 C 是 A 的子集，而且 $(n+1, 1, 2, \cdots, n, n+2, \cdots, 2n)$ 这个排列是 A 的元素而不是 C 的元素，因此 C 是 A 的真子集。这样为了证明 $N(A)>N(B)$，只要能证明 $N(C)=N(B)$ 就可以了，因为 $N(C)<N(A)$。

为此，下一步的工作就是要构造一个 B 与 C 之间的一一对应 f。考虑 B 的任一元素 $(y_1, y_2, \cdots, y_{2n})$，则 $|y_2-y_1|\neq n$。因此与 y_1 相差为 n 的数(这个数是惟一的)，一定是某一个 y_k，$k>2$。把 y_1 放到 y_k 的左边得一个新排列 $(y_2, \cdots, y_{k-1}, y_1, y_k, \cdots, y_{2n})$，这个排列一定是 C 的元素。作映射：

$$(y_1, y_2, \cdots, y_k, \cdots, y_{2n}) \to (y_2, \cdots, y_{k-1}, y_1, y_k, \cdots, y_{2n}),$$

不难证明这是 B 与 C 之间的一一对应。

2. 对称映射

有一个大家都很熟悉的故事，说的是德国数学家高斯在很小的时候，有一次老师给同学们出了一道算术题：

求 $1+2+3+\cdots+98+99+100$ 之和。

当高斯的小伙伴们还没有理清头绪的时候，高斯很快地说出了正确的答案。高斯求和的方法是：设

$S=1+2+3+\cdots+99+100$，

$S=100+99+98+\cdots+2+1$，

高斯把1和100相加，2和99相加，…，最后100和1相加，它们的和都是101，所以，$2S$ 的和是100个101，所以 $S=$

$\frac{100\times101}{2}=5050$．

高斯求和方法的妙处，就在于他将集合$A=\{1,2,\cdots,100\}$与A本身的元素之间进行了合理的配对，即建立了A到A的一一映射，这个映射使得每一对原象和象之和都等于101，从而把加法运算转化为乘法运算，大大地简化了求和的运算过程．在这里所用的映射，原象与象处于对称的位置，即与首、尾的距离相等．

推广这一思想，在数学解题中就形成一种对称思想，这里所说的对称，不单指几何意义上的对称，它泛指诸如形式、作用、地位等等方面的对称关系．利用对称思想解题的主要特点是，将问题中涉及的元素映射为它的对称元素，然后使原象与象配对以产生某种特殊的作用，如共轭、凑整、相消、互补等等，使之有助于问题的解决．因此，这里所说的对称思想本质上是一种配对思想．它与倍数映射不同的是，倍数映射只要求原象与象之间在数量上的配对，而对称映射则要求原象与象之间在某种性质上的配对．

例1 设a_1、a_2、$\cdots$、a_n为正实数，它们的和等于1，试证必有下列不等式成立：

$$\frac{a_1^2}{a_1+a_2}+\frac{a_2^2}{a_2+a_3}+\cdots+\frac{a_{n-1}^2}{a_{n-1}+a_n}+\frac{a_n^2}{a_n+a_1}\geqslant\frac{1}{2}.$$

这个问题自然可用平均值不等式、柯西不等式来证明，但明显地不等式的左边每一项都有对称性的项与之配对：

$$\frac{a_i^2}{a_i+a_{i+1}}\longrightarrow\frac{a_{i+1}^2}{a_i+a_{i+1}}.$$

因此，若令

$$x=\frac{a_1^2}{a_1+a_2}+\frac{a_2^2}{a_2+a_3}+\cdots+\frac{a_{n-1}^2}{a_{n-1}+a_n}+\frac{a_n^2}{a_n+a_1},$$

$$y=\frac{a_2^2}{a_1+a_2}+\frac{a_3^2}{a_2+a_3}+\cdots+\frac{a_n^2}{a_{n-1}+a_n}+\frac{a_1^2}{a_n+a_1}.$$

于是有

$$x-y=(a_1-a_2)+(a_2-a_3)+\cdots+(a_n-a_1)=0,$$

再利用$\frac{a_i^2+a_j^2}{a_i+a_j}\geqslant\frac{a_i+a_j}{2}$，即得：

$$\begin{aligned}x&=\frac{1}{2}[(x+y)+(x-y)]\\&=\frac{1}{2}(x+y)\\&\geqslant\frac{1}{4}[(a_1+a_2)+(a_2+a_3)+\cdots+(a_n+a_1)]\\&=\frac{1}{2}(a_1+a_2+\cdots+a_n)=\frac{1}{2}.\end{aligned}$$

例 2 如图6－4，设有一直角 MON，试在边 OM、ON 上及角的内部各求一点 A、B、C，使 $BC+CA=l$（l 为定长），且使四边形 $OACB$ 的面积最大.

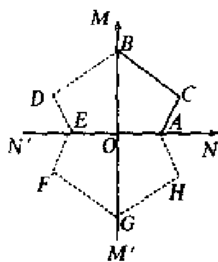


图 6－4

此题如果直接计算，无论是用三角方法还是解析方法都比较复杂．但是如果我们将直角 MON 连同四边形 $OACB$ 一起以 OM 为对称轴反射到$\angle MON'$，再以 NN' 为对称轴反射到下半平面，就得到一个八边形 $ACBDEFGH$，它是一个周长为定值 $4l$ 的八边形，由等周定理易知，当它为正八边形时面积最大．从而，这时四边形 $OACB$ 的面积也最大.

本题中连续使用了两次通过对称的配对，第一次使四边形 $OACB$ 与 $OEDB$ 配对，合成了五边形 $ACBDE$．第二次又使五边形 $ACBDE$ 与 $AHGFE$ 配对，合成八边形 $ACBDEFGH$.

例 3　求$[(\sqrt{3}+\sqrt{2})^{1980}]$的个位数字.

这是1980年欧洲四国的一道竞赛题，当年我国某出版社出版的一本书中介绍其解法时使用了数页篇幅，但如果利用对称思想则较易解决.

先去掉高斯符号：

$$(\sqrt{3}+\sqrt{2})^{1980}=(5+2\sqrt{6})^{990},$$

$(5+2\sqrt{6})^{990}$不是整数，加上一个$(5-2\sqrt{6})^{990}$就可以凑成整数，得到

$$(5+2\sqrt{6})^{990}+(5-2\sqrt{6})^{990}=A\in Z.$$

再注意到$0<5-2\sqrt{6}<1$，$0<(5-2\sqrt{6})^{990}<1$．即知所求的个位数字为$d\equiv A-1\pmod{10}$，$0\leqslant d\leqslant 9$．

在本题的解法中，实际上使用了一个巧妙的配对，即将$(5+2\sqrt{6})^{990}$的展开式中形如$5^m(2\sqrt{6})^{2n+1}=5^m\cdot 2^{2n+1}6\sqrt{6}$的项与$(5-2\sqrt{6})^{990}$的展开式中形如$-5^m\cdot 2^{2n+1}\cdot 6\sqrt{6}$的项配对，使它们相互抵消，留下来的所有各项都是整数，保证了其和A为整数，为解决问题提供了关键的便利.

3. 构形映射

在解数学问题时常常会碰到这样的情况，我们取某一元素a，就可以发现能在a的基础上构造出另一元素a'，a与a'在某一点上是紧密相关的，当我们把a映射为a'时，就揭露出了解题的关键．由于a'需要随机构造，不像对称元素那样是既定的，故我们把这种映射特称为构形映射．构形映射在本质上也是一种配对思想，即将a与以a为基础特别构造出来的另一元素a'配对.

例 1　用$P(n)$表示正整数n的分拆数，$P(n,k)$表示n的k－分拆数，$\frac{n}{2}\leqslant k\leqslant n$．证明：

$$P(n,k)=P(n-k),\ \frac{n}{2}\leqslant k\leqslant n.$$

例如，取 $n=5$，$k=3$，则 $n-k=5-3=2$．2的分拆有：$1+1$；2，故$P(2)=2$．而5的3-分拆有 $1+1+3$ 与 $1+2+2$，所以$P(5,3)=2$，即

$$P(5,3)=P(5-3).$$

现在记 n 的k-分拆的集合为 A，$n-k$ 的分拆构成的集合为B，那么，对于 A 中的任一元素

$$n=n_1+n_2+\cdots+n_k,\ (n_1\geqslant n_2\geqslant\cdots\geqslant n_k)$$

可与 B 中元素

$$n-k=(n_1-1)+(n_2-1)+\cdots+(n_k-1)$$

(其中有些项可以为0)配对．

剩下的问题只要证明，这是一一对应即可．

显然这个映射是单射，只要再证明它还是满射，即证明对 $n-k$ 的任一个分拆都对应 n 的一个 k-分拆．事实上，对 B 的任一分拆

$$n-k=m_1+m_2+\cdots+m_r,\ (m_1\geqslant m_2\geqslant\cdots\geqslant m_r)$$

因为 $n\leqslant 2k$，所以 $r\leqslant k$．那么 $n-k$ 的这个分拆可以与A 中的元素

$$n=\underbrace{(m_1+1)+(m_2+1)+\cdots+(m_r+1)}_{r个}+\underbrace{1+1+\cdots+1}_{k-r个}$$

对应，即在 A 中有原象，所以映射是满射．

在这个问题中，对 n 的任一个k-分拆都可同时构造出 $n-k$ 的一个分拆，将两者配对就找到了问题的关键．

例 2 某城市进行住房统计，如果用 c_k 表示住户不少于k人的住宅数($k=1$，2，…)，又将各个住宅里住户的人数排成 $d_1\geqslant d_2\geqslant\cdots$．证明

1) $c_1+c_2+\cdots=d_1+d_2+d_3+\cdots$

2) $d_1^2+d_2^2+\cdots=c_1+3c_2+5c_3+\cdots$

3) $c_1^2+c_2^2+\cdots=d_1+3d_2+5d_3+\cdots$

将每一个人用一个小黑点表示，则下图中每一纵列中的点数代表一所住宅中的人数：

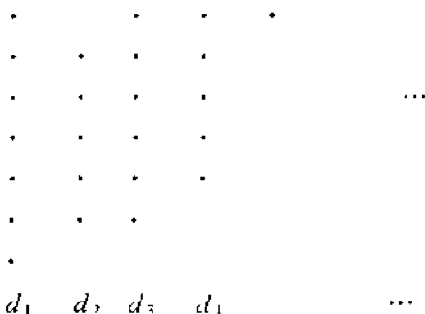


图6-5

图中黑点的总数就是

$$d_1+d_2+d_3+\cdots$$

但是另一方面，由上面这个点阵图同时派生出：第一行的点数是至少有1个人的住宅数，即 c_1，第二行的点数是至少有2人的住宅数，…，所以各行的总点数即图中黑点总数又为

$$c_1+c_2+c_3+\cdots$$

按纵列相加或按横行相加的结果应该是一致的，即有1)

$$c_1+c_2+c_3+\cdots=d_1+d_2+d_3+\cdots$$

为了证明2)，可在上图的基础上构造出一个新的点阵图：第一行的点不变，第二行的每点变成3点，第三行的每点变成5点，…最后一行(d_1行)的每点变成 $2d_1-1$ 点．

这时第一列的点数变为

$$1+3+5+\cdots+(2d_1-1)=d_1^2,$$

同理，第二列的点数为：

$$1+3+5+\cdots+(2d_1-1)=d_2^2.$$

余可类推，所以这时图中的总点数为

$$d_1^2+d_2^2+d_3^2+\cdots$$

另一方面，按横行计算点数时，第一行的点数仍为 c_1，第二行的点数是 $3c_2$，第三行的点数为 $5c_3$，…．总点数为

$$c_1+3c_2+5c_3+\cdots$$

总点数无论按横行或按纵列计算应该是不变的，故有2)成立．

类似地可证得3)．

d　d_2　d_3　d_4

图 6－6

在本题中，我们构造了点阵图6－5，使 c_i 与 d_i 配对，在图 6－5 的基础上再次构造了图6－6，使图 6－6 与图 6－5 配对，找到了问题的证明．

配对思想还有许多其他的表现形式，尽管表现的形式千差万别，但其基本思想是建立某一特定的映射，使得在这个映射之下，每一原象与它的象配成一对，得出在计算或论证中的某些方便．

三、划分思想

在数学论证中，经常要对一些数学对象分型划类．分类的方法有两种，一种是按照对象的内在属性进行逻辑分类，例如实数可分为有理数和无理数，三角形内角可分为锐角、直角和钝角．但是有些分类并不是按照数学对象本身的内在属性进行的，而是根据问题的需要临时规定一些分类标准进行人为的划分，以便分门别类进行论证或计算．对于后面这种分类，有很

多都是通过映射来实现的．所以我们把它归结为映射思想．

1．染色

有一道大家耳熟能详的数学趣题：

某展览馆有36间陈列室，每两间相邻陈列室之间有门可通，其入口与出口位置如图6－7所示．现在有人希望每间陈列室都参观到而且只经过每间展览室一次，这可能吗？如果可能，请为他设计一条参观路线；如不可能，请说明理由．

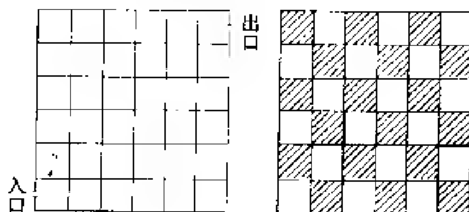


图6－7

本题的答案是不可能．因为如果像图6－7右边那样将36间展室依次相间地染上黑色和白色，则参观者无论怎样走法，从白色的展览室只能走到黑色的展览室，从黑色的展览室只能走到白色的展览室．任何人从白色的展览室进入，走了36间，只能走到黑色展室，如果不重复的话，决不能从另一白色的展室走出．因此，我们无法找到一条满足此人所要求的道路．

这个问题可以认为是“染色”的典型例题．

数学中有许多处于两种互相对立状态的对象，如正负、奇偶、正反、质数与合数等等，利用或制造两种对立的状态，把讨论的对象分为两类．为了表述的方便，用染色的办法来进行区分，以便对问题进行分析讨论，发现解题的线索．因此，染色的目的和作用是一种分类，在本质上是一种映射．

设集合 A 划分为两个集合 A_1 与 A_2（即 $A_1 \cup A_2 = A$，$A_1 \cap A_2 = \varnothing$），集合 B 由黑、白两种颜色组成 $B = \{黑、白\}$，二染色即为映射

$$\begin{cases} a_1 \longrightarrow 黑色; & a_1 \in A_1; \\ a_2 \longrightarrow 白色, & a_2 \in A_2. \end{cases}$$

把事物分成两类要用两种颜色，通常称为"二染色"；如果要把事物分成 k 类，则要用 k 种颜色，一般称为"k 染色"，染色方法在解决存在性问题中常常能发挥作用。

例 1 有 20 张卡片，今将数字从 0 至 9 的每一个都写在两张卡片上。试问，能否将这卡片排成一排，使得两个 0 相邻，两个 1 之间恰有 1 张卡片、两个 2 之间恰有 2 张卡片等等，直到两个 9 之间恰有 9 张卡片。（1965 年第 28 届莫斯科数学奥林匹克试题）

我们把 20 个位置依次染成白色和黑色：

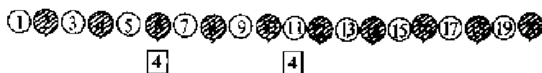


图 6－8

当我们从左至右放下某一个写有偶数(例如 4)的卡片时，要相隔 4 个位置再放第 2 张写有 4 的卡片，因而两个偶数卡片之间总相隔偶数个位置，因而它们所占住的位置必然一白一黑，因此5个偶数共占用 5 个黑位和 5 个白位。但是两个奇数卡片必须占据同色的位置，剩下的位置无论白色与黑色都是奇数，因此符合题设的排法是不存在的。

在这个问题中，有两个关键性的映射：一是用染色方法把 20 个位置划分为两类，二是将每一卡片 a 的位置 k 与另一张卡片 a 的位置 $k+a+1$ 进行配对。

例 2　设 S 是等边三角形 ABC 的三条边 AB、BC 和 CA 上所有点(包括顶点 A，B，C)的集合，对于把 S 分成两个不相重叠的子集的每一种分法，是否总能在分得的两个子集中至少找到一个子集，它含有某个直角三角形的 3 个顶点．证明你的结论．

设 S 已经划分为两个不相交子集 S_1，S_2，将 S_1 内的点染上红色，S_2 内的点染上蓝色，于是问题就转化为：

用红、蓝两种颜色任意去染等边三角形△ABC 周界上的点，问是否存在 3 个同色点，它们构成一个直角三角形的顶点．

答案是肯定的．如图6－9，设 D、E、F 分别为 AB，BC，CA 上的三等分点．D，E，F 中至少有两点同色，例如设 D，F 同为红色．若 AB 上还有红色点 G（$G\neq D$），则△FGD 符合所求；若 AB 上除 D 外都是蓝色点，则若 E 为蓝色，△EHB 符合所求；若 E 为红色，则当 C 为红色时，△CEF 符合所求；若 C 为蓝色，则作 $CK\perp AB$ 于 K，△CKB 符合所求．

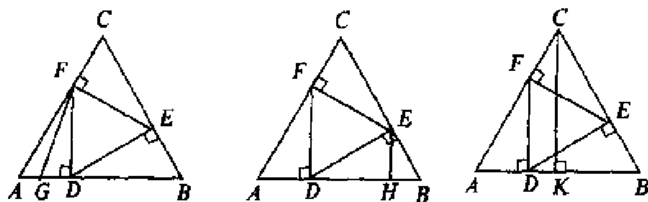


图6－9

2. 抽屉原理

《晏子春秋》里记载了一个"二桃杀三士"的故事：齐国的宰相晏婴想除掉齐景公的三名勇士，他深知用武力绝对制服不了

那三个人，便心生一计，以齐景公的名义赏赐三名勇士两个桃子，让他们按功劳的大小吃桃．由于三名勇士只有两个桃子，势必有两名勇士要合吃一个桃子，这两人感到羞愧而自杀了．剩下的一人也感到因争功吃桃而导致同伴自杀，实在对不起同伴也自杀了．晏子在这里运用了抽屉原理．

抽屉原理一般可以表述为以下三种类型：

(1) 把 $n+1$ 个元素任意放入 n 个集合中，那么至少有一个集合中的元素不少于 2 个．

(2) 把 $mn+1$ 个元素任意放入 n 个集合中，那么至少有一个集合中的元素不少于 $m+1$ 个．

(3) 把无限多个元素任意放入有限个集合中，那么至少有一个集合中有无限多个元素．

抽屉原理实质上也是一种映射：

设 A 为元素的集合，B 为由有限个集合为元素组成的集合，$B=\{C_1,C_2,\cdots,C_n\}$（C_i 为集合）．设 $a\in A$，将 a 放入集合 C_i 中相当于映射

$$a\longmapsto C_i,$$

把元素映射到 m 个集合也相当于把元素染上 m 种颜色．这个映射不是单射，也不一定是满射．它的特点是保证了集合 B 中存在某一元素 C_i，在 A 中的原象的个数不小于某一常数或有无穷多个．

抽屉原理虽然简单，但如果能巧妙地运用，有时可以收到出人意料的效果．

例 1 17 位科学家每一个都和其他人通信，在他们的书信往来中只讨论三个问题，而每两个科学家仅仅讨论其中一个问题．证明：至少有三个科学家，他们互相讨论同一个问题．（第 6 届 IMO 试题）

在17位科学家中任选一位A，其余的16位可以按与A讨论的题目的不同分为三类．那么这三类中必有一类（比方说与A讨论问题甲的一类）中至少有6个人．如果这6个人还有两个也互相讨论问题甲，则他俩与A三人都互相讨论同一问题甲，命题获证．

若这6人都不互相讨论问题甲，则他们互相之间只讨论问题乙和丙．在6人中再任选一人例如B，把其余的5人按与B讨论问题的不同分为两类．其中又必有一类（例如与B讨论问题乙的一类）不少于3人．若3人中还有两人互相讨论问题乙，则他俩与B三人互相讨论同一问题乙，命题获证．若这3人都不互相讨论问题乙，则他们都互相讨论问题丙，命题亦得证．

这一问题两次使用抽屉原理，用同一方法逐步论证，是运用抽屉原理的一个典型例题．

例2 已给一个集合，由任意10个互不相等的两位十进制正整数组成．求证：这个集合必有两个无公共元素的子集，它们所含各数之和是相等的．（第14届IMO试题）

10元素的集合子集的总数为$2^{10}=1024$个，除掉不应考虑的空集与全集还有1022个．另一方面，不相等的9个两位数的和不超过

$$99+98+\cdots+91=855,$$

这表明各子集中元素之和不会超过$855-9=846$种，因此至少有两个子集A、B，它们中元素之和相等．如果A与B中有公共元素，去掉其公共元素后，剩下的两个子集中的元素，其和亦相等．

从这一例题可以看到，使用抽屉原理解题时不在于具体论证之难，而在于设置"抽屉"之难．制造"抽屉"的常用思想方法，也离不开映射思想．

（1）分割图形．这种方法常用于解决这样一类问题：要证明若干元素能构成具有某种性质的图形，可把这些元素所在的集合划分为若干能产生所要求的图形的子集，把所有的元素映射到那些子集中去．

例 3 边长为 1 的正方形中任意放入 9 个点．证明：在以这些点为顶点的三角形中，必有一个三角形，它的面积不大于 $\frac{1}{8}$（如3点共线，则认为这个三角形的面积为 0）．

用对边中点的连线，把正方形分成 4 个面积为 $\frac{1}{4}$ 的小正方形．根据抽屉原理，至少有一个小正方形内有 3 个点，易证这 3 个点所成的三角形的面积小于小正方形的一半．

（2）利用剩余类．这种方法常用于解决与整数有关的问题，其方法是选择一个适当的模 m，将整数 a 映射到模 m 的剩余类：

$$a \longmapsto r, \qquad a \equiv r \pmod{m}.$$

特别地，若 $m=2$，则是按整数的奇偶性分类．

例 4 证明：存在形如 $11\cdots1$ 的 n 位数（$1 \leqslant n \leqslant 1999$），它能被 1999 整除．

考虑2000个这样的数：

$$1,\ 11,\ 111,\ \cdots,\ \underbrace{111\cdots1}_{2000\text{个}1}$$

根据抽屉原理，其中至少有两个用 1999 去除后的余数相同，从而二者之差必为 1999 的倍数．这个差具有形式

$$\underbrace{11\cdots1}_{j+k\text{个}} - \underbrace{11\cdots1}_{j\text{个}} = \underbrace{11\cdots1}_{k\text{个}} \times 10^{j}.$$

因为 $(1999, 10^{j}) = 1$，所以，$1999 \mid 11\cdots1$（k 个 1）．

抽屉原理不仅用于处理离散的量（集合中元素的个数），

推广其思想，又可用于处理连续的量．如关于面积的重叠原理：

重迭原理　假设平面上有 n 个区域，它们的面积分别是 S_1，S_2，…，S_n．如果把这 n 个区域按任何方式搬到一个面积为 S 的固定区域内部去，如果

$$S_1+S_2+\cdots+S_n>S,$$

那么，至少有两个区域具有公共点．

对于长度、体积也有类似的重叠原理．

用映射的语言来表述，则是：

将 $S_1, S_2, \cdots, S_n$（将面积为 S_i 的区域也记为 S_i）中的点，通过一个法则映射到区域 S 内，至少存在两个点 $A_i\in S_i$，$A_j\in S_j$，它们在 S 中有相同的象．

例5　设 P_1 是一个凸五边形，其5个顶点分别为 A_1，A_2，A_3，A_4，A_5，将 A_1 平移至 A_i（$i=2$、3、4、5）就得到另一凸五边形 P_i．证明：这5个凸五边形中至少有两个五边形有公共内点．

这是重叠原理的一个典型应用，我们只要找到一个凸多边形 P，使 P 包含5个凸五边形 P_1、P_2、P_3、P_4 和 P_5 于其内部，但 P 的面积却小于 P_1 面积 S 的5倍．

这是不难办到的，如图6－10，设 X 为 P_i 的任一内点，将 A_i 再平移回 A_1 时，X 平移到 Y，此时 $Y\in P_1$．在平行四边形 A_1A_iXY 内，A_1X 与 A_iY 有相同的中点 Z．因 P_1 为凸五边形，故 Z 也必为 P_1 的内点，且 $\overline{A_1X}:\overline{A_1Z}=2$．于是，若以 A_1 为中心，作位似变换，使位似比为2∶1，便得一五边形 $A_1B_2B_3B_4B_5$，令其为 P，则 P 符合所求．

3．同余

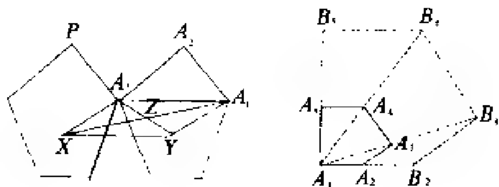


图6－10

在上节例4中，我们利用同余进行分类，通过分类来制造抽屉．同余是数学中一种非常重要的数学思想，取一个适当的正整数 m，把全体整数按其用 m 去除所得余数的不同分为 m 个类，余数为 i $(0\leqslant i\leqslant m)$的一类记作 M_i，称为模 m 的剩余类．全体整数的一个划分：

$$Z=\{M_0,\ M_1,\ \cdots,\ M_{m-1}\},$$

称为模 m 的一个划分，从 M_i $(0\leqslant i\leqslant m-1)$各取一个数为代表组成的集合称为模 m 的一个完全剩余系，简称完系．

整数的许多性质对它的剩余类具有“遗传性”，在处理这类问题时，可以把必须对全体整数讨论的问题，转化为对其一个剩余类的研究，常常可以大大地简化论证和计算．

同余也是一种映射，对每一个整数 $i\in Z$，作映射

$$i\xrightarrow{f}M_i.$$

这种映射显然应属于划分思想．

数论中的许多重大成果都是利用同余思想取得的．初等数学中也常借助同余思想来简化解题中的论证和计算．

例1 当4444^{4444}写成十进制时，它的各位数字之和是 A，A 的各位数字之和是B，求 B 的各位数字之和C．（第17届IMO试题）

这是一个与进位制有关的问题．它基于这样一条定理：

正整数 n 与它的各位数字之和模 9 同余．因此我们有 $4444^{4444} \equiv A \equiv B \equiv C \pmod{9}$．

下一步就只要实际算出 $4444^{4444} \equiv 7 \pmod{9}$，并估计出 $0 < C \leqslant 13$．从而推出 $C = 7$．

例 2 求所有的正整数 m、n，使得

$$1! + 2! + \cdots + m! = n^2.$$

这是一个不定方程，它的一个基本思想，是利用方程的两边对某些模数必须同余，由此排除不合格的正整数，从而大大缩少要研究的范围．

首先注意到，当 $m \geqslant 5$ 时，

$$1! + 2! + 3! + 4! + 5! + \cdots$$
$$\equiv 1! + 2! + 3! + 4!$$
$$\equiv 3 \pmod{5}.$$

但是，对任意的正整数 n，都有 $n^2 \equiv 0, 1, 4 \pmod{5}$，这说明当 $m \geqslant 5$ 时，方程没有解．剩下的问题就是对 $m = 1, 2, 3, 4$ 直接检验方程有无解的问题了．不难验证，方程有两解 $(1, 1)$ 和 $(3, 3)$．

例 3 试确定具有下述性质的所有正整数 n，集合

$$M = \{n, n+1, n+2, n+3, n+4, n+5\}$$

可以分成两个不相交的非空子集，使得一个子集中所有元素的积等于另一个子集的所有元素之积．（第 12 届 IMO 试题）

由于 M 的元素是 6 个连续整数，故可利用剩余类来研究．在 M 中补充一个 $n+6$，则

$$n, n+1, n+2, n+3, n+4, n+5, n+6$$

构成模 7 的一个完系．于是，我们就可以在模 7 上来讨论 n 的性质．

若 $n + 6 \not\equiv 0 \pmod{7}$，则有 $a_i \in M$，使 $a_i \equiv 0 \pmod{7}$．

从而在 M 的两个子集中，一个子集中所有数的积模 7 为 0，另一个子集中所有数的积模 7 不为 0，这是不符合题目要求的。

若 $n+6\equiv 0\pmod 7$，则令 M 的两个子集中的元素之积均为 m，则有

$$m^2=n(n+1)(n+2)(n+3)(n+4)(n+5)$$
$$\equiv 1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6=720\equiv -1\pmod 7.$$

但 $m^2\equiv 1, 4, 2\pmod 7$，矛盾。

综上所述，满足题设条件的 n 不存在。

在这一节中，我们简略地论述了通过适当的映射来进行分型划类的方法，这种分类并不是借助于那些一望而知的与元素本身内在属性直接有关的方法进行的，而是通过精心设计的巧妙映射来实现的。染色是数学中划分思想的形象说法，而同余则是划分思想的一种具体应用。

四、赋值思想

赋值是这样一种数学思想方法，在处理一些没有明显的数量关系的对象时，为了论证的需要，将每一对象按一定的规律映射为一个特殊的数值，然后通过研究这些数值的个数、分布规律，或者进行某些运算，使题中隐蔽的条件和关系明朗化，从中发现规律，以期找到解决问题的途径。

1. 布尔代数

数学中在处理两种相互对立的数学对象时，常常用两种颜色将它们着色，以便分型划类。但是在初等数学中，着色虽能起到区分的作用，但没有定义颜色的运算，所以，我们有时将 1 和 0 两个数字来代替两种颜色，用赋值 1 或 0 来代替染色。因为1和 0 这两个数字可以建立二进制数系统，也可以建

立起布尔代数系统，借助于代数运算更有利于发现解题的关键。

布尔代数(这里只指二值代数)，也称为开关代数，它只有两个数值 0 和 1，它有加法和乘法，加法与乘法运算的结果如下：

+	0	1
0	0	1
1	1	1

×	0	1
0	0	0
1	0	1

众所周知，电子器件的电路、计算机的软件都是利用布尔代数的运算来编制的。

在初等数学中也可借助于布尔代数的运算来解决某些数学问题：将要考虑的数学对象中的一部分 a 映射为 1，另一部分 b 映射为 0．

$$a \longmapsto 1, \qquad b \longmapsto 0.$$

然后借助于布尔代数的运算求解问题．

例 1 一条长为 n 的线段分成了 n 段，两端的端点染成蓝色，其余分点任意染成红色或蓝色．证明：端点染色不同的线段必有偶数条．

如图 6－11，我们把两个端点同色的线段称为甲种线段，两个端点异色的线段称为乙种线段．并把红色端点与0对应，蓝色端点与 1 对应．

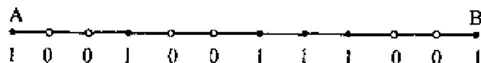


图 6－11

那么一个甲种线段的两个端点之和为 0，而一个乙种线段

的两个端点之和为1．因为除A、B两点外，每个端点都作为两条相邻线段的端点，在加法运算中使用两次，其和均为0，故所有线段端点之和为$1+0+\cdots+0+1=0$．但另一方面，两个乙种线段端点之和为1．因此，若乙种线段为奇数条的话，则两两求和后其和为0，剩下最后一个乙种线段其端点之和为1．与已证明的所有线段端点之和为0的结果矛盾．

例2 有两个大小不同的同心圆盘，均分成$2n$个相同的小扇形．外盘固定，内盘可以绕两圆的公共中心转动，将内、外两盘的所有小扇形都染成红、蓝两种颜色之一，且染同一颜色的小扇形两盘中总计为$2n$个．证明：可将内盘转到一个适当位置使两盘中扇形对齐，而对应颜色不同的扇形不少于n对．

把红色扇形对应于0，蓝色扇形对应于1，那么当对应的两个扇形同色时，其和为0，对应的两个扇形异色时，其和为1．因此，用a_1、a_2，…，a_{2n}和b_1、b_2，…，b_{2n}分别记外盘和内盘各扇形的值．当固定外盘、内盘转动到b_i与a_1相对时，对应的扇形之和组成一个$2n$维的布尔向量

$$(a_1+b_i,\ a_2+b_{i+1},\ \cdots,\ a_{2n}+b_{i-1}) \qquad ①$$

本题的结论便转化为证明：

存在一个正整数$i(1\leqslant i\leqslant 2n)$使得①式中的$2n$个分量中，1的个数不少于$n$个．

假定内盘中红色扇形有k个$(0\leqslant k\leqslant 2n)$，则内、外两盘中两种颜色扇形的个数如下表所示：

内盘	红色k个	蓝色$2n-k$个
外盘	红色$2n-k$个	蓝色k个

内盘中每个红色扇形与外盘中每个蓝色扇形构成一个1，共k^2

个1．$2n-k$内盘蓝色扇形与外盘$2n-k$个红色扇形构成$(2n-k)^2$个1．所以，当外盘固定，内盘转动$2n$次绕过一周后，共有

$$(2n-k)^2+k^2=4n^2-4nk+2k^2=2(2n^2-2nk+k^2)$$
$$=2[n^2+(n-k)^2]\geqslant 2n^2.$$

由于在$2n$次转动后至少要出现$2n^2$次1，故至少有一次要出现n个1．

在这两个例题中，我们可以看到，染色虽然也有划分两种对立状态的作用，但没有赋值，就很难进一步推证．

2．特殊值法

解选择题时我们常常使用一种被称为"特殊值法"的方法．例如，要解下面一道问题：

当a，b是两个不相等的正数时，下列三个代数式：

甲：$\left(a+\dfrac{1}{a}\right)\left(b+\dfrac{1}{b}\right)$；乙$\left(\sqrt{ab}+\dfrac{1}{\sqrt{ab}}\right)^2$；

丙$\left(\dfrac{a+b}{2}+\dfrac{2}{a+b}\right)^2$中间，值最大的一个是（　　）．

（A）必定是甲；　　（B）必定是乙；

（C）必定是丙；　　（D）不能确定，与a，b的值有关．

取$a=1$，$b=2$，可排除(B)，(C)；再取$a=2$，$b=3$，可排除(A)，故应选(D)．

当我们需要处理的数学对象并非处于两种对立状态，而是呈多元状态时，当然不再适宜于赋值1或0来求解．这时可以具体问题具体分析，根据问题的需要赋予一些数学对象以特殊的数值．一般地，赋值的目的仍是把不能具体运算的对象转化为数值运算来寻求证明．

80年代初期，某省有一道饶有趣味的数学竞赛试题：

7只茶杯口全朝下，每次操作可将其中4只翻转，问能否经过有限次操作，将茶杯全变成口朝上。

有一个巧妙的方法解此题，我们令

杯口朝上⟶1，杯口朝下⟼ -1．

开始时7只杯口全朝下，其对应值的乘积为

$$(-1)\times(-1)\times(-1)\times(-1)\times(-1)\times(-1)\times(-1)=(-1)^7=-1.$$

每次翻转4只茶杯，相当于改变上式中左边7个因子中4个的符号，并不影响其乘积的符号，因此，不管操作多少次，乘积结果永远是 -1．而7只杯口都朝上时，乘积应为 $1^7=1$．这意味着，要通过所说的操作使得7个杯口皆朝上是不可能的．

例1 组装甲、乙、丙三种产品，需用 A，B，C 三种零件．每件甲产品需要 A，B 各2个；每件乙产品需用 B，C 各1个；每件丙产品需用2个 A 和1个 C．用库存的 A，B，C 三种零件组成 p 件甲产品，q 件乙产品和 r 件丙产品，则剩下2个 A 零件和1个 B 零件，但 C 零件恰好用完．试证：无论如何改变甲、乙、丙的件数，也不能把 A，B，C 三种零件都恰好用完．

我们给 A，B，C 三种零件规定一个价格：

$A\longmapsto 1$ 元，$B\longmapsto 2$ 元，$C\longmapsto 4$ 元．

则甲、乙、丙三种产品的价格分别为：

甲产品⟶ $2\times1+2\times2=6$，

乙产品⟶ $1\times2+1\times4=6$，

丙产品⟶ $2\times1+1\times4=6$，

剩余零件⟶ $2\times1+1\times2=4$．

所以，零件的总价格为

$$6(p+q+r)+4.$$

因此，不论怎样改变甲、乙、丙三种产品的数量结构，都不能使总价格恰为6的倍数，也就不能把全部零件用完。

本题如用列方程求解则比较麻烦，用赋值的办法则比较简便。

例2 有男、女孩各 n 人($n\geqslant 3$)围坐成一个圆圈，若顺序相邻的3人中，恰有一个男孩的有 x 组，恰有一个女孩的有 y 组，那么 $x-y$ 一定能被3整除。

用 a_i 表示第 i 个孩子（$i=1,2,\cdots,n$），给每个孩子对应一个数：男孩对应 $w=-\dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i$，女孩对应 $w^2=-\dfrac{1}{2}-\dfrac{\sqrt{3}}{2}i$，$w^3=1$，于是，在每个3人组中：

$$a_ia_{i+1}a_{i+2}=\begin{cases}1, & \text{当 } a_i,\ a_{i+1},\ a_{i+2} \text{ 为三男或三女时}, \\ w, & \text{当 } a_i,\ a_{i+1},\ a_{i+2} \text{ 为二男一女时}, \\ w^2, & \text{当 } a_i,\ a_{i+1},\ a_{i+2} \text{ 为一男二女时}.\end{cases}$$

把所有3人组对应的数相乘，其积为

$$(w^2)^x w^y=w^{2x}w^y=w^{3x}w^{-x}w^y=w^{y-x}.$$

另一方面，在所有3人组中，每一个 a_i 都出现恰好3次，所以乘积又应该是

$$a_1^3a_2^3\cdots a_n^3=1.$$

于是 $w^{y-x}=1$，从而 $y-x$ 是3的倍数。

3. 排序思想

在一些数学问题中涉及到一批可用数量来刻画的对象（如数的大小、线段的长短、角的大小等等），如果它们之间没有顺序，杂乱无章，往往会使许多有利于解题的条件隐蔽起来，给解题带来困难。因此，在解题之前，可能的话给它们规定一种次序，天然的或人为的，都常常有助于解题的思考。

排列是一种特殊的赋值映射：

设 a 是被研究的元素，把它的顺号排在 n，则相当于

$$a \longmapsto n .$$

排序思想有助于问题解决，可从下例看出：

某片树林中所有的树高度都不超过20米，任何两株树之间的距离，都不超过它们的高度之差．那么我们一定可以用一条不超过40米的篱笆把所有的树都围起来．

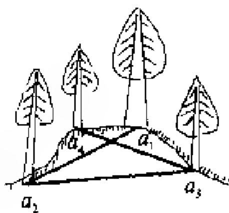


图 6－12

问题的条件好像不太充分，这些树生长的状态没有作充分交待．它们杂乱无章，“远近高低各不同”，篱笆似乎无从围起．

但我们可借助排序思想来解此题：

假设这片树林中共有 n 株树，把这些树从高到矮排一个顺序，设其高度依次是 a_1，a_2，…，a_n．根据题意应有

$$a_1 > a_2 > \cdots > a_n .$$

现在按这个次序把树的根部依次连接起来(如图6－12)，就得到一条折线．这条折线的长度不大于

$$(a_1 - a_2) + (a_2 - a_3) + \cdots + (a_{n-1} - a_n)$$
$$= a_1 - a_n < a_1 \leqslant 20(\text{米}).$$

于是，我们只要沿着这条折线把它的两侧都围起来，所有的树便都围进去了，而篱笆的长度是折线的两倍，不超过40米．

解决这道问题时，把所有的树按其高度排序起了关键的作用，它揭示了篱笆应遵循的走向，即根部的那条折线，同时保证了那条折线的长度不超过树的最大高度．由此可见排序原理的妙用．

例 1　n 个正数 $a_1, a_2, \cdots, a_n$ 的算术平均数 $A = \frac{a_1 + a_2 + \cdots + a_n}{n}$ 与几何平均数 $G = \sqrt[n]{a_1 a_2 \cdots a_n}$ 之间具有关系：

$$G \leqslant A. \tag{①}$$

这是一个著名的重要不等式，它有各种各样的证法，包括数学归纳法．但奇怪的是，长期以来，使用数学归纳法的证明都是使用的间接归纳法，一直到 20 世纪中叶才有人利用排序思想给出一些直接的归纳法证明．

不妨假定 $a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n$，那么

$$a_1 \leqslant A \leqslant a_n.$$

因此　$(A - a_1) \geqslant 0, (a_n - A) \geqslant 0,$

那么　$(A - a_1)(a_n - A) = Aa_n + Aa_1 - A^2 - a_1 a_n \geqslant 0,$

或者　$A(a_1 + a_n - A) \geqslant a_1 a_n$.

假定对于任意 $n-1$ 个非负实数，①式成立．考虑 $n-1$ 个实数

$$a_2, a_3, \cdots a_{n-1}, (a_1 + a_n - A),$$

则　$$\frac{a_2 + a_3 + \cdots + a_{n-1} + (a_1 + a_n - A)}{n-1} = \frac{nA - A}{n-1} = A,$$

$$\sqrt[n-1]{a_2 a_3 \cdots a_{n-1}(a_1 + a_n - A)} \geqslant \sqrt[n-1]{\frac{a_1 a_2 \cdots a_n}{A}}.$$

由归纳假定

$$A \geqslant \sqrt[n-1]{\frac{a_1 a_2 \cdots a_n}{A}},$$

所以　$A^{n-1} \geqslant \frac{a_1 a_2 \cdots a_n}{A}$,

或　$A^n \geqslant a_1 a_2 \cdots a_n$.

即得①式对 n 也成立．

在许多不等式的证明中，更多地用到排序思想，因为它基

于下面的重要定理：

设 a_1，a_2，…，a_n；b_1，b_2，…，b_n 是 $2n$ 个非零实数，满足条件

$$a_1\geqslant a_2\geqslant\cdots\geqslant a_n,\quad b_1\geqslant b_2\geqslant b_3\geqslant\cdots\geqslant b_n .$$

则 $$a_1b_1+a_2b_2+\cdots+a_nb_n\geqslant a_1b_{i_1}+a_2b_{i_2}+\cdots+a_nb_{i_n}$$
$$\geqslant a_1b_n+a_2b_{n-1}+\cdots+a_nb_1 . \qquad ②$$

许多重要的不等式都可以通过排序不等式②推导出来。

例 2 设 x_1，x_2，…，x_n 都是正整数，且满足

$$x_1+x_2+\cdots+x_n=x_1x_2\cdots x_n , \qquad ①$$

求 x_1，x_2，…x_n 中的最大值。

由于条件的对称性，不妨设

$$x_1\leqslant x_2\leqslant\cdots\leqslant x_n . \qquad ②$$

由于这一排序，就改变了条件的对称性，相当于增加了一个附加条件

$$x_{n-1}\geqslant 2\ (n\geqslant 2) .$$

否则，若 $x_{n-1}=1$，则 $x_1=x_2=\cdots=x_{n-1}=1$，代入①将有

$$(n-1)+x_n=x_n .$$

这是矛盾的。有了这一附加条件，就可推得

$$x_n=\frac{x_1+x_2+\cdots+x_{n-1}}{x_1x_2\cdots x_{n-1}-1}$$
$$\leqslant\frac{x_1x_2\cdots x_{n-2}+x_1x_2\cdots x_{n-2}+\cdots+x_1x_2\cdots x_{n-2}x_{n-1}}{x_1x_2\cdots x_{n-1}-1}$$
$$=\frac{(n-2+x_{n-1})x_1x_2\cdots x_{n-2}}{x_1x_2\cdots x_{n-1}-1}$$
$$\leqslant\frac{(n-2+x_{n-1})x_1x_2\cdots x_{n-2}}{x_1x_2\cdots x_{n-1}-x_1x_2\cdots x_{n-2}}$$
$$=\frac{n-2+x_{n-1}}{x_{n-1}-1}=1+\frac{n-1}{x_{n-1}-1}\leqslant n .$$

当 $x_1=x_2=\cdots=x_{n-2}=1$，$x_{n-1}=2$ 时，x_n 确能取得最大值 n．也就是 x_1，x_2，…，x_n 中的最大值．

在这一节中，我们讨论了几种赋值思想，赋值的特点是：将一些数学对象映射为一些特殊选定的数值，然后借助于这些特殊的运算来推理．赋值的本质是一种映射思想，把数学对象映射为特殊的数值，由于这些数值的引入，增加了解题的条件，因而能为解题提供有益的帮助．

五、表示思想

表示是数学中最常用的术语．

一个数学元素常常有不同的表示方法．例如，每一个正整数 m，在十进制中可以惟一地表示成

$$a_n\cdot 10^n+a_{n-1}\cdot 10^{n-1}+\cdots+a_1\cdot 10+a_0$$

的形式，其中 $a_i\in\{0, 1, 2, \cdots, 9\}$，$a_n\neq 0$，$i=0, 1, 2, \cdots, n$．除此之外，正整数还有二进制、三进制等表示方法，还有标准分解表示法等等．

解析几何的一个基本思想就是把点用坐标的形式表示，用坐标表示的方法本身也是多种多样的，如直角坐标、极坐标、面积坐标等等．

表示是一种特殊的映射，它把元素 a 从一种形式映射为另一种表示形式：

$a\longmapsto a$ 的另一种表示形式．

正是数学元素的不同表示形式，大大地丰富了数学的内容，提供了解题的技巧．因此，善于运用不同的表示是一种重要的数学思想方法．

1．正整数的表示

如前所述，正整数的表示方法是多种多样的．哪怕是最简

单的1，它在数学中有哪些表示，恐怕任何人都难以全面回答，例如，1可表示为：

$$1=\operatorname{tg}\frac{\pi}{4}=\sin^2\theta+\cos^2\theta,$$

$$1=e^0=\log_a a=\lg 10,\ (a>0)$$

$$1=C_n^0=C_n^n,$$

$$1=\omega^3=\left(-\frac{1}{2}+\frac{\sqrt{3}}{2}i\right)^3,$$

$$1=\lim_{x\to 0}\frac{\sin x}{x}=\lim_{x\to 0}\frac{e^x-1}{x},$$

$$1=\int_0^{\frac{\pi}{2}}\sin x\,dx,$$

$$1=(-1)\times(-1)\times\cdots\times(-1)=(-1)^{2k},$$

1是模 m 的一个剩余类，

1是“大衍求一”中的1，

1是一件工程的总体，

1是必然事件的概率，

等等，简直可以无穷无尽地列举下去.

因此，在解数学问题时，决不能把“1”简单地只当作最小的正整数，那样将会使你寸步难移．相反的，如能对“1”这个数字的多种表示进行这样或那样的联想，将会得到奇特的效果．

例1 在区间 $0\leqslant x\leqslant 2\pi$ 中，求出所有的实数 x，使满足

$$2\cos x\leqslant\left|\sqrt{1+\sin 2x}-\sqrt{1-\sin 2x}\right|\leqslant\sqrt{2}.$$

如在根号下将1表示为 $\sin^2 x+\cos^2 x$，则被开方式化为完全平方，从而有

$$\left|\sqrt{1+\sin 2x}-\sqrt{1-\sin 2x}\right|=\left|\,|\sin x+\cos x|-|\sin x-\cos x|\,\right|,$$

下一步便可以分区间讨论了.

有些数常常可转化用三角函数表示，从而可简化问题．如

求$\sqrt{2}\cdot\sqrt{2+\sqrt{2}}\cdot\cdots\cdot\underbrace{\sqrt{2+\sqrt{2+\sqrt{2+\cdots+\sqrt{2}}}}}_{n\text{重根号}}$．

如果注意到$\cos45^\circ=\dfrac{\sqrt{2}}{2}$，$\cos\dfrac{45^\circ}{2}=\sqrt{\dfrac{1+\cos45^\circ}{2}}=\dfrac{\sqrt{2+\sqrt{2}}}{2}$，于是猜想

$$\underbrace{\sqrt{2+\sqrt{2+\cdots+\sqrt{2}}}}_{n\text{重根号}}=2\cos\frac{45^\circ}{2^{n-1}}.$$

用数学归纳法极易证明．

例 2　试证：对任意的正整数n，数

$$A=1+\frac{1}{2}+\cdots+\frac{1}{n}$$

都不是整数．

本题如用数学归纳法去证明是不恰当的，其关键是利用正整数的一种表示方式，令

$$i=2^{m_i}t_i,\quad i=1,\ 2,\ \cdots,\ n.$$

其中m_i为非负整数，t_i为奇数．在m_1，m_2，…，m_n诸数中，必有一个最大者，而且是惟一的，设为m_s．再令$t_1t_2\cdots t_n=T$．现在设

$$A=1+\frac{1}{2}+\cdots+\frac{1}{n}.$$

用$2^{m_s-1}T$乘A的两边，若A为整数，则左边$2^{m_s-1}T\cdot A$为整数，而右边除$\dfrac{1}{2^{m_s}t_s}$这一项外，在乘以$2^{m_s-1}T$之后也均为整数，但$2^{m_s-1}T\cdot\dfrac{1}{2^{m_s}t_s}=\dfrac{t_1\cdots t_{s-1}t_{s+1}\cdots t_n}{2}$，分子为奇数而分母为2，故不能为整数，从而右边不为整数．这个矛盾证明了所

要的结论．

由以上二例可见，在解与整数有关的问题时，要充分利用数的各种表示形式．一个正整数有多种表示形式，常见的有：

（1）带余形式：$a=km+r$；

（2）标准分解式：$n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_k^{\alpha_k}$（$p_1<p_2<\cdots<p_k$ 为质数，α_1，α_2，$\cdots\alpha_k$ 为正整数）；

（3）$n=2^m t$ 的形式：m 为非负整数，t 为奇数；

（4）各种不同的进位制，如二进制、三进制、十进制等等．

合理地选择整数的表示形式，常对解题大有帮助．

2．坐标方法

笛卡儿的坐标法构建起了沟通数与形的桥梁．点与坐标之间的互相映射、互相表示成了学习数学必须具备的基础知识和基本技能．在阐述数形结合思想时，我们还将详细讨论，故此处从略，只略举一例，以备一格而已．

例 1 若 $|a|<1$，$|b|<1$，求证

$|ab\pm\sqrt{(1-a^2)(1-b^2)}|\leqslant 1$．

从要证的不等式左边的结构使我们联想到几何中的托列迷定理：

圆内接四边形两组对边之积的和小于两对角线的积．

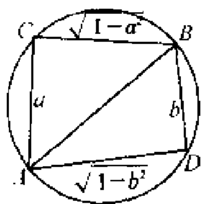


图 6－13

如图 6－13，以 $AB=1$ 为直径作圆，并在 AB 两侧分别取圆上的两点 C 和 D，使 $AC=a$，$BD=b$，那么就得到映射

$a\longmapsto AC$，

$b\longmapsto BD$，

$\sqrt{1-a^2}\longmapsto BC$,

$\sqrt{1-b^2}\longmapsto AD$,

$ab+\sqrt{1-a^2}\sqrt{1-b^2}\longmapsto AC\cdot BD+AD\cdot BC$.

上式的右边恰好就是托列迷定理的左边，注意到 $AB=1$，$CD\leqslant 1$，便有

$AC\cdot BD+AD\cdot BC\leqslant AB\cdot CD\leqslant AB=1$.

这就证明了 $\left|ab+\sqrt{1-a^2}\sqrt{1-b^2}\right|\leqslant 1$.

例 2 设 E 为$\triangle ABC$ 的高 CD 上任一点，直线 AE 交 BC 于 F，直线 BE 交 AC 于 G．求证直线 DF、DG 关于 CD 对称．

如图 6－14，要证 DF、DG 关于 CD 对称，即要证$\angle BDF=\angle ADG$，或$\angle ADG$ 与$\angle ADF$ 互补，也即证直线 DG 与直线 DF 的斜率互为相反数．

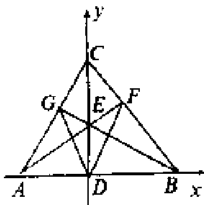


图 6－14

以 AB、DC 为坐标轴建立如图 6－14 之坐标系，即可将各点映射为它的坐标：

$A\longmapsto(a,\ 0)$,

$B\longmapsto(b,\ 0)$,

$C\longmapsto(0,\ c)$,

$E\longmapsto(0,\ k)$,

$D\longmapsto(0,\ 0)$.

于是，AC，BE 的方程分别为：

$cx+ay=ac$,

$kx+by=bk$.

由方程组解出 G 点的坐标为

$$x_G=\frac{ab(k-c)}{ak-bc},\ y_G=\frac{ck(a-b)}{ak-bc}.$$

从而得 DG 的斜率为

$$k_1=\frac{ck(b-a)}{ab(k-c)}.$$

类似地，可得 DF 的斜率为

$$k_2=\frac{ck(a-b)}{ab(k-c)}.$$

于是有 $k_1=-k_2$.

3. 图表方法

有些需要解决的数学问题，涉及的对象比较庞杂，讨论起来十分困难，有时可设计一张图或列一个表，使讨论的对象比较直观，一目了然，这样有助于问题的解决.

设计图表是一种特殊的映射方法，它把一些数量关系转化为图表，但在建立其映射关系时，往往需要很强的构造能力，所以要与构造思想很好地结合起来.

例 1 马路上有 6 个车站，今有一辆汽车由第一站驶向第六站. 沿途各站都可以自由上下乘客，但此辆汽车在任何时刻都只能载客 5 人. 试证明：在此 6 站中必定有两对(4 个)不同的车站 A_1，B_1；A_2，B_2（A_1 在 B_1 之前，A_2 在 B_2 之前），使得没有乘客在 A_1 站上车而在 B_1 站下车，也没有乘客在 A_2 站上车而在 B_2 站下车.

如果用 a_{ij}（$1\leqslant i<j\leqslant 6$）表示从第 i 站上车，而在第 j 站下车的旅客数，它满足 $0\leqslant a_{ij}\leqslant 5$，就可以把各站上下车人数列成一个表，如图 6－15 所示.

显然在表的右上角有一个 3×3 的块，在这个块中的 9 个数所表示的上车人数，在第四站前还没有一个下车. 所以9个数之和不能超过 5，又由于它们都是非负数，必然至少有 4 个

0．这4个0分布在三行三列上，必然至少有两个0既不同行又不同列，设为 a_{mn} 和 a_{ij}（$m<n$，$i<j$，$m\neq i$，$n\neq j$），那么就没有旅客在第 m 站上车而在第 n 站下车、也没有旅客在 i 站上车而在第 j 站下车．

上 \ 下	一	二	三	四	五	六
一		a_{12}	a_{13}	a_{14}	a_{15}	a_{16}
二			a_{23}	a_{24}	a_{25}	a_{26}
三				a_{34}	a_{35}	a_{36}
四					a_{45}	a_{46}
五						a_{56}
六						

图 6－15

本题的列表解法实质上是建立了一个十分巧妙的映射，将在第 i 站上第 j 站下的人数，映射到上三角阵第 i 行第 j 列的位置，从而构造出一个有用的 3×3 矩阵．

例 2 剧场出售票 $2n$ 张，每张票价 5 元，每人限购一张．开始时售票处没有钱，而排队购票的前 $2n$ 人中，一半人持有5元币，另一半人持有 10 元币．要让售票处不发生找钱困难，这 $2n$ 个购票人有多少种不同的排队方法．

这是一个计数问题，但它的计算太复杂，我们设计一个与之同构的图形：在平面直角坐标系 Oxy 中，用原点 O 表示售票处，购票者依次以横坐标为 1，2，…，$2n$ 的点表示．令每个持10元币者对应 $+1$，持 5 元币者对应 -1．显然，任何前 k

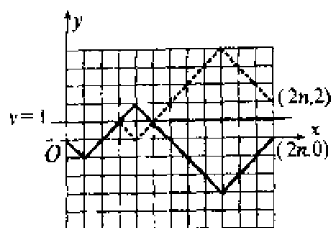


图 6－16

个购票者所对应的数之和必须小于1，否则就会出现找不出钱的情况．今以 k 为横坐标，前 k 个人所对应的数之和为纵坐标，就对应平面上的一个格点．顺次连结这些点就得到从(0，0)到(2n，0)的一条折线，这条折线与 $2n$ 个人的排队方法有一一对应的关系，并且不发生找钱困难的排队方法所对应的折线不与直线 $y=1$ 相交．

现在的问题是转化为计数从(0，0)到(2n，0)之间的折线，这种折线共有 n 段上升，n 段下降，每一条都由它上升的 n 段的位置决定，故有折线 C_{2n}^{n} 条．而与直线 $y=1$ 相交的折线数为 C_{2n}^{n-1} 条（将这样的折线在第一个交点处以 $y=1$ 为轴向上反转，就得到一条从(0，0)到(2n，2)的折线，它有 $n+1$ 段上升，$n-1$ 段下降，故有 C_{2n}^{n-1} 条），从而不发生找钱困难所对应的折线有 $C_{2n}^{n}-C_{2n}^{n-1}=\frac{1}{n}C_{2n}^{n-1}$ 条．

本题中综合使用了赋值、映射、构形等多种方法，需要较高的技巧．

六、同态思想

现代数学研究抽象的运算．让我们来分析一下，运算的本

质是什么？考虑正整数的加法运算：

$$3+5=8,$$

我们把它改写为

$$(3,\ 5)\longmapsto 8.$$

因为3，5，8都是正整数集N的元素，(3，5)是N与N的一个序偶．或者说是笛卡儿积$N\times N$的一个元素．因此，我们可以说：正整数的加法是集合N的笛卡儿积$N\times N$到集合N中的一个映射．

据此，"运算"可以抽象地定义为：

设A和B是两个集合，A的笛卡儿积$A\times A$到集合B的一个映射，称为集合A的一个二元运算．

这个定义可表示为图6－17．

$$a\cdot b=c.$$

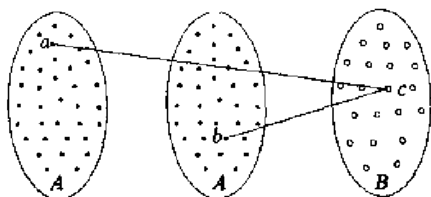


图6－17

当B是A的子集时，则称运算封闭．

如果集X中有一种运算"·"，集Y中有一种运算"＊"，f是集X到Y的映射，如果在映射中能保持运算不变，即对于X中的任意两个元素x_1，x_2，若

$$x_1\longmapsto f(x_1)=y_1,\ x_2\longmapsto f(x_2)=y_2,$$

便有 $x_1\cdot x_2\longmapsto f(x_1)*f(x_2)=y_1*y_2.$

我们就说f是一个"从X到Y的(关于运算"·"与"＊"的)

同态”．如果 f 又是一一对应，则称为“同构”．

例如，设 X 为全体整数，$Y=\{0, 1, 2, \cdots, m-1\}$ 是模 m 的一个完系，f 是同余映射：

$$km+r \longmapsto r \quad (k\in X,\ r\in Y).$$

易知，对于 X 中任意两个元素 k_1m+r_1，k_2m+r_2，有

$$(k_1m+r_1)+(k_2m+r_2)\longmapsto r_1+r_2,$$

因此，f 是 X 到 Y 的关于加法“+”的同态，并且易知 f 对于乘法“·”也是同态，因为

$$(k_1m+r_1)(k_2m+r_2)\longmapsto r_1\cdot r_2.$$

我们知道，复数有多种表示形式，如：

代数式：$a+bi$；

三角式：$r(\cos\theta+i\sin\theta)$；

指数式：$re^{i\theta}$；

向量式：$\overrightarrow{OA}$．

这些表示之间都可以建立一一对应的关系，它们都是同构映射．

例如复数的代数式与三角式之间的对应：

$$a+bi\longleftrightarrow r(\cos\theta+i\sin\theta),$$

其中 $r=\sqrt{a^2+b^2}$，$\cos\theta=\frac{a}{r}$，$\sin\theta=\frac{b}{r}$．

对于复数：

$$a_1+b_1i\longleftrightarrow r_1(\cos\theta_1+i\sin\theta_1),$$

$$a_2+b_2i\longleftrightarrow r_2(\cos\theta_2+i\sin\theta_2),$$

则有

$$\begin{aligned}&r_1(\cos\theta_1+i\sin\theta_1)+r_2(\cos\theta_2+i\sin\theta_2)\\&=(r_1\cos\theta_1+r_2\cos\theta_2)+i(r_1\sin\theta_1+r_2\sin\theta_2)\\&\longleftrightarrow(a_1+a_2)+i(b_1+b_2)=(a_1+b_1i)+(a_2+b_2i),\end{aligned}$$

$$r_1(\cos\theta_1+i\sin\theta_1)\cdot r_2(\cos\theta_2+i\sin\theta_2)$$
$$=r_1r_2(\cos\theta_1\cos\theta_2+i\cos\theta_1\sin\theta_2+i\sin\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2)$$
$$=r_1r_2(\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2)+ir_1r_2(\sin\theta_1\cos\theta_2+\cos\theta_1\sin\theta_2)$$
$$\longleftrightarrow(a_1a_2-b_1b_2)+i(a_1b_2+a_2b_1)$$
$$=(a_1+ib_1)(a_2+ib_2).$$

这就证明了复数的代数式与三角式同构.

利用同态或同构映射解题，在很多场合比一般的映射更为有效. 我们在前面讨论的各种映射中，就有不少是同态或同构的.

下面我们看几个利用同态思想解题的例.

1. 同态

例 1 $2n$ 个篮球队参加循环赛，每两队之间均需赛一场，共需比赛$C_{2n}^2=\dfrac{2n(2n-1)}{2}=n(2n-1)$场. 但为了保证各队的休息，每队每天至多进行一场比赛. 由于每队需要进行$2n-1$场比赛，因此至少要赛$2n-1$天. 现在的问题是：$2n-1$天是否一定能够赛完？如果可能，怎样安排出比赛的日程表？

答案是肯定的. 具体安排日程表的办法是利用"同余".

先考虑前$2n-1$个队，即$1,2,\cdots,2n-1$. 若其中两个队i，j满足

$$i+j\equiv k(\bmod 2n-1),$$

则在第k天安排i与j比赛.

在这样的安排下，前$2n-1$个队每队每天至多赛一场. 事实上，在第k天，i的对手为$k-i$或$2n-1+k-i$,

$$\begin{cases} i\longmapsto k-i, & \text{当 } k>i; \\ i\longmapsto (2n-1)+k-i, & \text{当 } k\leqslant i. \end{cases}$$

当且仅当

$$i\equiv k-i \pmod{2n-1}$$

时，i 在第 k 天没有安排进去.

将上式变形为 $2i\equiv k \pmod{2n-1}$，再乘以 n，得

$$2ni\equiv nk \pmod{2n-1},$$

或 $$i\equiv nk \pmod{2n-1}.$$

所以，在第 1，2，3，…，$2n-1$ 天没有安排进来的队分别为 n，1，$n+1$，…，$2n-1$.

对于每个 i，$k-i \pmod 7$ 互不相同（$k=1, 2, \cdots, 2n-1$），所以 i 与其他 $2n-2$ 个队都比赛了一场.

最后让第 $2n$ 队安排进来，让它和当天没有安排比赛的队比赛，即第 k 天与 $nk \pmod{2n-1}$ 比赛.

如果是 $2n-1$ 个队，也需要 $2n-1$ 天才能赛完. 因为第一天有一个队轮空，这个队以后要用 $2n-2$ 天进行 $2n-2$ 场比赛，所以比赛的天数不少于 $2n-1$ 天. 另一方面 $2n-1$ 个队可采用 $2n$ 个队的安排方式安排比赛，其中有几场轮空. 所以 $2n-1$ 天也是足够的.

在这个问题的解决中，利用了“同余”映射，也就是利用了“同态”映射.

例 2 在第四章中，我们介绍了“抓火柴”游戏的取胜策略. 它是利用映射到二进数来解决的. 设三堆火柴的根数分别是 28，19，13 支，将它们映射为二进数：

28⟶011100，

19⟶010011，

45⟶101101 .

这是对加法运算的同态映射.

现在我们来考虑另一个数学游戏.

如图 6－18，在 8×8 的棋盘中放有三枚棋子，两人轮流移动棋子，每一次可将任何一枚棋子移动若干格，不准不移、也不准同时移动两枚以上的棋子，移动的方向必须按箭头运行的方向，可以几枚棋中落在同一格中，所有的棋子都要移到画有圆圈的格子里，谁将最后一枚棋子移到这个画圆圈的格子里，谁就取得胜利．问谁有取胜策略．

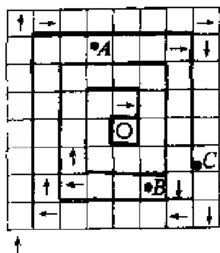


图 6－18

这个问题虽然形式上与抓火柴的游戏不同，但实际上是一样的，我们只要作映射：

一枚棋子移到画圆圈方格的步数⟶一堆火柴数⟶一枚棋子移动的步数⟶在一堆火柴中取走的根数．

于是：　　棋子 $A\longrightarrow 28\longrightarrow 011100$，

棋子 $B\longrightarrow 19\longrightarrow 010011$，

棋子 $C\longrightarrow 45\longrightarrow 101101$．

这显然是一个同态映射．

这是一个“非正则数组”，与取火柴的游戏一样，先取者必胜．如果是“正则组”，则应让对手先走，后取者有必胜策略．

2．类比

我国南北朝时的唯物论学者范缜曾经写过一本叫做《神灭论》的著作，他在论证人死亡后并没有灵魂时写道：“神之于形，犹利之于刀，未闻刀没而利存，生容形亡而神在．”范缜在这里用的是一种类比推理，他将“形”与“刀”对应、“神”与“利”对应，并且将它们之间的依存关系也互相对应：

神⟼利，

形⟼刀，

形亡$\longmapsto$刀没，

神在$\longmapsto$利存．

在这个对应中，不仅建立了元素之间的对应，还建立它们之间的关系对应，与同态映射极为相似．当然范缜所用的对应，并不是严格意义下的数学中的同态映射，只不过与它有相似之处，可以进行类比而已．所以，我们把这一节的标题称为类比．

类比是一切科学研究的重要手段之一，当然也是数学研究的重要手段之一．

用类比研究数学的思路可表示为：事物 A 有性质 a，b，c，$d\cdots$，事物 B 有与之相应的性质 a'，b'，c'，$d'\cdots$，从 A 的性质之中可导出新的性质 x，从而估计从 B 的性质可推出新的性质 x' 与之相应．

我们可以将类比思想与同态映射思想作如下的类比：

同　态	类　比
集合 A 有元素 a，b，c，$d\cdots$	事物 A 有性质 a，b，c，$d\cdots$
集合 B 有元素 a'，b'，c'，$d'\cdots$	事物 B 有性质 a'，b'，c'，$d'\cdots$
A 与 B 之间有映射关系： $a\longrightarrow a'$， $b\longrightarrow b'$， …	A 与 B 之间有“相似”关系： $a\longrightarrow a'$， $b\longrightarrow b'$， …
A 中某些元素通过 A 的运算得出新元素 x，B 中这些元素的象通过 B 的运算得出新元素 x'，则 $x\longrightarrow x'$．	A 中某些性质推出新的性质 x，B 中相应的性质大概也可推出新的性质 x'，使 $x\longrightarrow x'$．

由对比可以看出，对于同态映射，有 $x\rightarrow x'$ 的结论是肯定的；对于类比，有 $x\rightarrow x'$ 只是一种“大概可以”，结论并不可

靠，更不能代替数学的证明.

类比虽然只是一种猜测，未必正确，但是波利亚说得好：“如果把这种猜测的似真当作肯定性，那将是愚蠢的；但是忽视这种似真的猜测将是同样愚蠢甚至更为愚蠢的.”

波利亚在他的名著《怎样解题》一书中，在谈到类比时，曾经引用“求均匀四面体的重心”这个问题作为例子．从三角形的重心通过类比推出四面体的重心的性质．他还用下面的类比来推导这一同一结论：

点是零维的，线段是一维的，平面是二维的，空间是三维的.

线段的边界是两个零维元素(端点)，而其内部是一维的；

三角形有 3 个零维边界元素(顶点)和 3 个一维边界元素(三边)，而其内部是二维的；

四面体有 4 个零维边界元素(顶点)、6 个一维边界元素(棱)和 4 个二维边界元素(面)，而其内部是三维的.

把这些数字列表如下：

	零维	一维	二维	三维
线段	2	1		
三角形	3	3	1	
四面体	4	6	4	1

①

① 2 1

① 3 3 1

① 4 6 4 1

细心的读者不难发现这些数字是杨辉三角的一部分，这预示在线段、三角形、四面体之间有一种值得注意的规律性.

既然线段的重心与其两端点的重心相重合，三角形的重心与其三顶点的重心相重合，

为什么我们不能设想四面体的重心与其四顶点的重心相重合呢？并且，线段的重心按1∶1的比例划分其两端点的距离．三角形的重心按2∶1的比例来划分任何顶点与其对边中点的距离．为什么我们不应该猜想四面体的重心是按3∶1的比例来划分其任何一个顶点与其对面重心间的距离呢？

波利亚在评论这一类比推理时写道："说上述问题所提出的猜测是错误的，说这种美妙的规律性竟遭破坏，这点总叫人觉得极不可能．认为和谐的简单秩序不会骗人这样一种感觉，在数学及其他科学领域中指引着作出发现的人们，并表达为拉丁格言：简单是真理的标志！"

例1 设在桌面上有一个丝线做成的线圈，它的周长是$2l$，我们又用纸片剪成一个直径是l的圆形纸片．证明：不管线圈作成什么样的曲线，我们都可以用此圆形纸片完全盖住它．

本题的难点在于线圈可以是任意的形状的曲线．如果线圈是一个特殊的形状，例如圆，问题就好办得多，只要像图6－19(a)那样，把纸片的圆心重合于线圈的圆心就可以了．

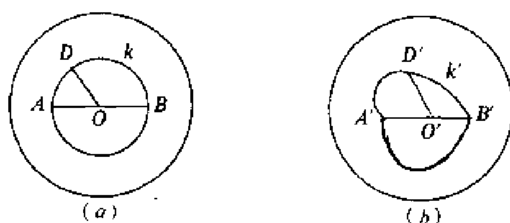


图6－19

k既然是一个圆，圆的主要特征是它的圆心和直径，取一条直径AB，A、B两点把圆分成相等的两部分．k'为任意的曲线，也有把k'分成相等的两部分的两点A'、B'，线段$A'B'$

有中点 O'．于是可这样作出类比：

$$O \longmapsto O',$$
$$A \longmapsto A',$$
$$B \longmapsto B'.$$

$$\underbrace{\begin{pmatrix}k\text{ 上任一点 }D\text{ 与 }O\\ \text{的距离}\leqslant l;\end{pmatrix}}_{\text{已有的}} \longrightarrow \underbrace{\begin{pmatrix}k'\text{ 上任一点与 }O'\\ \text{的距离}\leqslant l.\end{pmatrix}}_{\text{要证的}}$$

从而得到解题的方法．

例 2　如图6－20，已知四面体 $ABCD$ 中，$AB=CD$，$AC=BD$，$AD=BC$．证明：它的四个面都是锐角三角形．

三角形与三面角有密切的联系，我们考虑下面的类比：

三角形		三面角
①三角形的三边；	⟶	①三面角的三个面角；
②三角形的任意两边之和大于第三边；	⟶	②三面角的任意两个面角之和大于第三面角；
③欲证三角形的每边都小于 l，可证三边之和$\leqslant 2l$．	⟶	③欲证每一个面角小于 θ，可证三个面角之和$\leqslant 2\theta$．

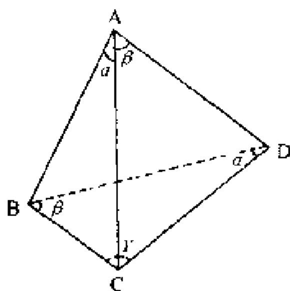


图 6－20

根据这一类比，很快找出证明本题方法．利用 $AB=CD$，$AC=BD$，$AD=BC$ 可推出四面体的 4 个面都是全等三角形．三面角 $A-BCD$ 的三个面角 α，β，γ 恰好是 $\triangle BCD$ 的三个内角．因为 $\alpha+\beta+\gamma=180°$，所以，α，β，γ 的每一个小于 $\frac{180°}{2}=90°$．

§6.2 数形结合思想

在这一节中，我们将讨论一种综合运用转化与构造的思想，即数形结合的思想．数形结合把形转化为数或把数转化为形，首先它是一种转化思想．但是另一方面，无论是形转化为数或数转化为形，都要进行必要的构造，构造出适当的形或数以供转化．所以，数形结合的思想是综合运用转化与构造的典型思想．

初等数学研究的对象大体上可分为"数"与"形"，它们研究的内容和使用的方法虽然有区别，但其间并没有不可逾越的鸿沟，它们互相联系，互相渗透，互相转化．笛卡儿创立的解析几何，建立了形数结合的典范．

一般地说，几何图形比较"直观"，代数问题比较"抽象"．抽象的代数问题，一旦与几何图形结合，就往往易于估测其结果，找出论证的方法；几何中的难题，一旦化为代数问题，也往往有一定的运算方法和步骤可循，因而易于求解．因此，形数结合的思想就成为初等数学中一种极为重要的思想，在解决复杂问题时，尽量把数与形结合起来，使它们互相转化，使一些复杂的问题变得简单，隐晦的问题变得明显，有利于问题的解决．

一、数形转化

数形结合的思想早在毕达哥拉斯时代就萌芽了．毕达哥拉斯学派对数与形的关系有特殊的理解，他们把单位 1 想像为一个点，由点的各种不同排列可以组合成各种图形，而各种不同的图形就与相应的数对应．这个学派关于许多数的性质的发现，都是以数形结合的方法为出发点得出的．

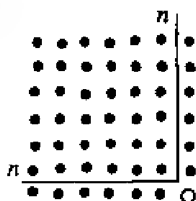


图 6－21

例如，正方形数是由 n^2 个点组合成的方阵，如果再加上 $2n+1$ 个点，就得到由 $(n+1)^2$ 个点组合成的方阵，如图6－21所示．即有公式

$$n^2+2n+1=(n+1)^2 .$$

如果取 $2n+1=(2m+1)^2$，就会得到

$$(2m+1)^2+(2m^2+2m)^2=(2m^2+2m+1)^2,$$

这正好是毕达哥拉斯定理的形式．

正是通过图形与数字的联系的思考，毕达哥拉斯学派形成了"万物皆数"的世界观．他们认为既然数字是一些点的组合，而且是不可分的单位，因此，只有能表示成整数或整数之比的数才是合理的，否则就是不合理的．但是具有讽刺意义的是，他们发现有不能表示为整数或整数之比的量，也是通过几何图形而发现的．

无论是中国或者是希腊、阿拉伯等国家最初都是用图形来解二次方程的．例如古代阿拉伯数学家是这样解二次方程

$$x^2+10x=39 \text{ 的：}$$

在图 6－22 中，AB 表示未知数 x，作正方形 $ABCD$，延长 DA 到 H，DC 到 F，使 $AH=CF=5$，以 DH 为边作正方

形，则Ⅰ，Ⅱ，Ⅲ的面积分别为 x^2，$5x$，$5x$．三者之和即方程的右边39．39加上面积Ⅳ，即得 $39+25=64=HD^2$，从而 $HD=8$，$x=AD=HD-5=8-5=3$．这种图解法大抵相当于配方法，但只取了正根．

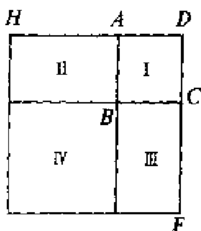


图 6－22

形数结合思想发展到了笛卡儿时代，就发生了质的飞跃．笛卡儿曾提出了一种解决各类问题的万能的模式：

（1）把任何问题转化为数学问题；

（2）把任何一个数学问题化为一个代数问题；

（3）把任何一个代数问题归结为求解一个方程式．

他的计划过于庞大，自然无法实现，但却在解决几何问题中得到了成功的运用．根据他的模式，他创立了解析几何学，开辟了数形结合的新纪元．

我国古代数学家十分注意数形结合，例如刘徽在《九章》注的原序中说："又所析理以辞，解体用图，庶几约而能周，通而不黩，览之者思过半矣．"所以他在注《九章》中始终注意以数显形和以形验数．

数形结合可以帮助我们加深对某些数学问题的理解．例如古希腊时代的几何作图三大难题千百年来在几何中无法解决，只有化为代数问题之后，他山之石，可以攻玉，才充分揭露问题的本质，使这一难题获得彻底的解决．

其次，数形结合思想对启迪解题思路，简化解题方法的作用更是极为有效的．

例如：如图 6－23，设$\angle XOY$ 为定角，P 为其角平分线上的定点，对 P 任作一直线交角的两边于 A、B．求证：$\frac{1}{OA}$

$+\frac{1}{OB}=$定值.

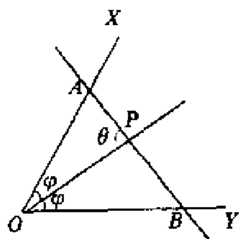


图6-23

这个问题利用几何方法直接证明需要一定的技巧，而且要作补助线. 但若利用计算则较为容易：

由正弦定理：设$\angle APO=\theta$，$\angle AOB=2\varphi, OP=a$，则

$$\frac{OA}{\sin\theta}=\frac{a}{\sin(\varphi+\theta)},$$

$$\frac{1}{OA}=\frac{\sin(\varphi+\theta)}{a\sin\theta};$$

$$\frac{OB}{\sin(\pi-\theta)}=\frac{a}{\sin(\theta-\varphi)},\quad \text{即}\frac{1}{OB}=\frac{\sin(\theta-\varphi)}{a\sin\theta}.$$

所以 $$\frac{1}{OA}+\frac{1}{OB}=\frac{\sin(\theta+\varphi)+\sin(\theta-\varphi)}{a\sin\theta}=\frac{2\sin\theta\sin\varphi}{a\sin\theta}$$

$$=\frac{2\sin\varphi}{a}=\text{定值}.$$

又例如，已知$0<y<x<\frac{\pi}{2}$，$\mathrm{tg}x=n\,\mathrm{tg}y$，求$u=x-y$的最大值.

这是一个二元函数的条件极值，且涉及三角和代数，似乎很不简单，但利用几何图形，则相当简便. 如图6-24，设$\triangle ABC$中，$\angle A=x$，$\angle B=y$，$\angle B<\angle A<\frac{\pi}{2}$，$CD\perp AB$于$D$，且$AD=1$，$DB=n$，则$\mathrm{tg}x=n\,\mathrm{tg}y$.

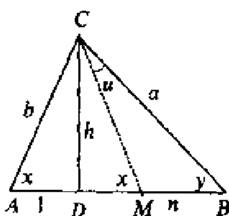


图6-24

在AB上取点M，使$AD=DM$，所以$\angle CMA=x$，$u=x-y$.

因为 $S_{\triangle ABC}=\frac{1}{2}ab\sin\angle ACB$，$S_{\triangle MCB}=\frac{1}{2}qb\sin u$，又因为 $S_{\triangle MCB}=\frac{1}{2}CD\cdot MB=\frac{1}{2}(n-1)\cdot CD$，$S_{\triangle ABC}=\frac{1}{2}CD\cdot(n+1)$，所以 $\sin u=\frac{2S_{\triangle MCB}}{ab}=\frac{2S_{\triangle ABC}}{ab}\cdot\frac{n-1}{n+1}=\frac{n-1}{n+1}\cdot\sin\angle ACB\leqslant\frac{n-1}{n+1}$，从而 $u\leqslant\arcsin\frac{n-1}{n+1}$.

以上诸例虽小，但足以说明数形结合的作用.

以下各节，我们将分别讨论初等数学中几种典型的数形结合思想.

二、以数表形和以形验数

1．以数表形

以数表形的基本思想是把几何图形的证明转化为代数方法的运算，常用的方法有坐标法、三角法、复数法或向量法等.

如前所述，笛卡儿的把一切问题代数化的计划虽然无法实现，但他的计划在解决欧氏几何的问题上却获得了圆满的成功．从理论上讲，欧氏几何的命题都可以用坐标法转化为代数问题来解决.

一道陌生的几何题摆在面前，常常使人感到无从下手，但一经代数化之后，情况就不同了．几何的证明转化成了一系列代数的运算，运算也许十分繁杂，但有既定的目标和固定的方法，总可以一步一步地算下去，最后得出所要的结论．特别是电子计算机被普遍地采用以来，这一作用更为明显.

例 1 圆内接四边形 $ABCD$ 中，对角线 AC 与 BD 垂直相交于 Q，P 为外接圆的圆心，E，F 分别为 BC，AD 的中点．求证：$QE-PF$.

在图形中有 $DB\perp AC$，故可取它们为坐标轴建立如图 6－25 的坐标系.

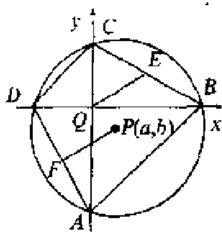


图 6－25

设圆心 P 的坐标为 $(a,\ b)$，半径为 r，就可求出圆方程:

$$(x-a)^2+(y-b)^2=r^2,$$

再分别令 $x=0$，$y=0$，可算出四边形各顶点的坐标:

$$A\ (0,\ b-\sqrt{r^2-a^2}),$$

$$B\ (a+\sqrt{r^2-b^2},\ 0),$$

$$C\ (0,\ b+\sqrt{r^2-a^2}),$$

$$D\ (a-\sqrt{r^2-b^2},\ 0).$$

再根据中点公式，定出 E、F 的坐标:

$$E\left(\frac{B_x}{2},\ \frac{C_y}{2}\right),\ F\left(\frac{D_x}{2},\ \frac{A_y}{2}\right).$$

(其中 B_x 表示 B 点的横坐标，余类推).再根据距离公式又可算出:

$$QE=\frac{1}{2}\sqrt{B_x{}^2+C_y{}^2},\ PF=\frac{1}{2}\sqrt{B_x{}^2+C_y{}^2}.$$

于是证得 $QE=PF$.

证明过程虽然略显繁复，但思路却至为简单，只要反复地运用勾股定理以及 QE、PF 用 a，b，r 所表示的代数式.据此，不难把它对应地"翻译"为综合法的证明.分析 P，Q，E，F 的坐标，就不难发现:

$$E_x+F_x=P_x+Q_x,\ E_y+F_y=P_y+Q_y.$$

从而知 EF 与 PQ 有相同的中点，这意味着 $PEQF$ 为平行四边形.由此得到启发，应先从证明 $PEQF$ 为平行四边形入手.图中有许多圆周角、直角可以利用，利用内错角或同位角相等的

办法即可证明 $QE \parallel PF$，$QF \parallel PE$.

从而证明四边形 $QPFE$ 为一个平行四边形，于是有 $QE=PF$.

例 2 求证：三角形外心、重心、垂心三点共线.

要证明三点共线，只需建立适当的坐标系，再计算由坐标组成的一个三阶行列式为 0 即可.但对一个任意三角形，在直角坐标系中，计算其外心、垂心并不容易，题中又有重心，故宜选用重心坐标.

在平面上任取一个三角形 $A_1A_2A_3$，作为坐标三角形. 对于平面上任一点 M，下述三个三角形面积的比值

$$\triangle MA_2A_3 : \triangle MA_3A_1 : \triangle MA_1A_2 = m_1 : m_2 : m_3,$$

叫做点 M 的重心坐标，也叫面积坐标，记为

$$M(m_1 : m_2 : m_3).$$

经过不太复杂的计算知，对于任意的 $\triangle A_1A_2A_3$，其三顶点的重心坐标为

$$A_1=(1,\ 0,\ 0),\ A_2=(0,\ 1,\ 0),\ A_3=(0,\ 0,\ 1) .$$

重心 G、外心 O、垂心 H 的坐标分别为

$$G=(1,\ 1,\ 1),\ O=(\sin 2A_1,\ \sin 2A_2,\ \sin 2A_3),$$
$$H=(\operatorname{tg}A_1,\ \operatorname{tg}A_2,\ \operatorname{tg}A_3) .$$

于是 G、O、H 共线的条件是

$$\begin{vmatrix} 1 & 1 & 1 \\ \sin 2A_1 & \sin 2A_2 & \sin 2A_3 \\ \operatorname{tg}A_1 & \operatorname{tg}A_2 & \operatorname{tg}A_3 \end{vmatrix} = 0 .$$

在几何证明中，经常的大量的工作是研究线段或角之间的关系，包括位置关系和度量关系，位置关系又常常转化为度量关系来处理.角的大小也可以转化为研究它们的三角函数值的

增减问题.因此，许多几何题要证的结论，常可归结为一些由线段通过四则运算组成的等式或不等式.当我们把那些线段归结到一个三角形或几个三角形内后，就可通过正弦定理或余弦定理把它们用三角函数表示出来，转化为三角函数的计算，而三角函数的计算是比较有规则可循的.

例 3 （托列迷定理）如图 6－26，设 $ABCD$ 为圆内接四边形，证明：$AB \cdot CD + AD \cdot BC = AC \cdot BD$.

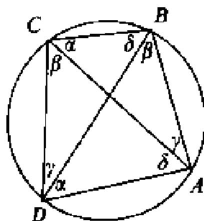


图 6－26

在第五章中我们曾讨论过这一问题，那是从分析线段再构造相似三角形的角度出发的.现在再从数形结合的角度讨论一下.

本题中的 6 条线段都是同圆的弦，把它们表示成三角函数后，有许多相等的角可供利用，有利于简化三角函数的计算.而命题的结论又正是典型的线段四则等式，故可用三角方法通过计算来求证.

设圆的半径 $R=1$，则有

$AB=2\sin\alpha, BC=2\sin\gamma, DA=2\sin\beta$,

$CD=2\sin(\alpha+\beta+\gamma), AC=2\sin(\alpha+\gamma)$,

$BD=2\sin(\alpha+\beta)$,

利用和差化积与积化和差公式极易导出结论.

复数的几何表示把坐标、三角、向量熔于一炉，平面上的点与复数有一一对应的关系，因此，把几何问题转化为复数问题来处理是很自然的事情.

例 4 如图6－27，在任意△ABC 的三边上向外作等边三角形.求证：这三个等边三角形的中心构成一个等边三角形.

这是一个涉及正三角形的问题，自然可考虑用 1 的三次方

根．为此，先介绍下面的定理：

设有三点 A，B，C，且正$\triangle ABC$的周界按字母 A，B，C 的顺序为逆时针方向，则 $A + B\omega + C\omega^2 = 0$．（1，$\omega$，$\omega^2$ 为 1 的三次单位根）

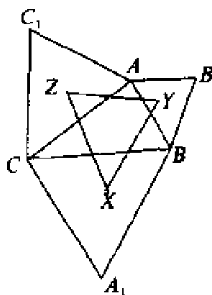


图 6－27

如图 6－28，设 D 是 BC 的延长线上一点，且 $BC = CD$．则由复数乘法

$$A - C = (D - C)\omega = (C - B)\omega,$$

即　$A + B\omega - C(1 + \omega) = 0$．

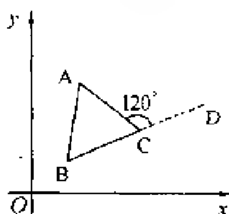


图 6－28

因为　$1 + \omega + \omega^2 = 0$，所以 $A + B\omega + C\omega^2 = 0$．

现在回到原题，因

$$X = \frac{1}{3}(A_1 + B + C),$$

$$Y = \frac{1}{3}(A + B_1 + B),$$

$$Z = \frac{1}{3}(A + C + C_1).$$

从而

$$\begin{aligned}3(X + Y\omega + Z\omega^2) &= (A_1 + B + C) + (A + B_1 + B)\omega \\ &\quad + (A + C + C_1)\omega^2 \\ &= (A_1 + B\omega + C\omega^2) + (B + B_1\omega + A\omega^2) \\ &\quad + (C + A\omega + C_1\omega^2) = 0\text{（利用定理）}.\end{aligned}$$

所以，$X + Y\omega + Z\omega^2 = 0$，显然上述定理的逆命题也正确，故$\triangle XYZ$为正三角形．

2．以形验数

“以形验数”基于这样的思想：

建立坐标系以后，许多代数式都具有特定的几何意义，例

如，若 a，b 是两个正实数，则二次根式$\sqrt{a^2+b^2}$表示以 a，b 为直角边的直角三角形的斜边之长．代数中的函数和方程可表示为平面上的一条曲线或空间中的一个曲面．利用这些联系，可以构造出某些几何图形，把代数式转化为几何图形，再根据几何图形的性质讨论代数问题．这一思想方法很好地把转化与构造两个基本数学思想融为一体了．

例 1 试确定方程组

$$\begin{cases} x+y+z=3, & ① \\ x^2+y^2+z^2=3, & ② \\ x^3+y^3+z^3=3 & ③ \end{cases}$$

的所有实数解．

这是 70 年代美国数学奥林匹克的一道试题．在当年发表的一些解答中都是利用 x，y，z 的对称函数化为一元三次方程来求解的，那种解法计算很复杂且需要相当的技巧．

如果利用数形结合的思想，考虑到①可改写为

$$x+y+z-3=0, \qquad ④$$

它表示空间中的一个平面，在 Ox，Oy，Oz 轴上的截距均为 3，与坐标原点 O 的距离为

$$d=\frac{|1\times0+1\times0+1\times0-3|}{\sqrt{1^2+1^2+1^2}}=\sqrt{3}.$$

而方程②是一个以原点为中心，$\sqrt{3}$为半径的球面，它与平面④相切，切点为(1，1，1)．因此方程只有惟一的实数解(1，1，1)．由于它的解已由方程①和②完全决定，方程③实际上是“多余”的，只要检验(1，1，1)是否适合方程③就可以了．

在这一思想的启发下，我们可以立即写出这一问题的解答：

将方程②减去方程①的 2 倍，得

$$x^2+y^2+z^2-2x-2y-2z=-3,$$

或 $$(x-1)^2+(y-1)^2+(z-1)^2=0.$$

因为 x，y，z 均为实数，立得 $x=1$，$y=1$，$z=1$.

例 2 解方程 $x+\frac{x}{\sqrt{x^2-1}}=\frac{35}{12}$.

如果将方程两边平方，将变为四次方程，求解甚繁．因 $x^2>1$，$|x|>1$，且显然 $x>0$．故可在原方程中作三角代换．$x=\sec\theta$，原方程化为：

$$\frac{1}{\cos\theta}+\frac{1}{\sin\theta}=\frac{35}{12}.$$

如果再注意到 $\frac{35}{12}=\frac{5}{3}+\frac{5}{4}$.

则可知 $x=\frac{5}{3}$ 或 $\frac{5}{4}$.

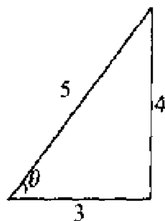


图 6－29

这可由图 6－29 的直角三角形推出．

函数或方程可以表示为几何图象，通过对几何图象的观察，可直观地发现许多问题，简化推理和计算．所以，"以形验数"的另一重要方面是观察函数的几何图象，把代数问题转化几何问题．

例 3 已知一元二次方程 $7x^2-(k+13)x+k^2-k-2=0$ 有两个实根．问 k 为何值时，方程的两根分别在 0 与 1 和 1 与 2 之间．

函数 $y=7x^2-(k+13)x+k^2-k-2$ 的图象与 x 轴有两个交点，而且依题意这两个交点在 0 与 1 和 1 与 2 之间．观察其图象(图 6－30)：

$f(0)>0, f(1)<0, f(2)>0$.

故可得两个不等式：

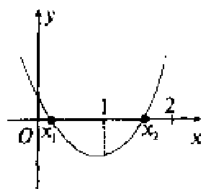


图 6－30

$$\frac{f(0)}{f(1)}<0 \text{ 和 } \frac{f(1)}{f(2)}<0 .$$

于是问题转化为解关于 k 的不等式：

$$\frac{f(0)}{f(1)}=\frac{(k+1)(k+2)}{(k+2)(k-4)}<0 \text{ 和 } \frac{f(1)}{f(2)}=\frac{(k+2)(k-4)}{k(k-3)}<0 .$$

例4　若 $a>b>c>0$，比较 $c^a\cdot b^c\cdot a^b$ 与 $c^b\cdot b^a\cdot a^c$ 的大小．

这是一个可用排序原理来研究的不等式问题．但考虑对数函数的曲线 $y=\ln x$，如图 6－31，在 x 轴上取 $A(a, 0)$，$B(b, 0)$，$C(c, 0)$．并过 A，B，C 分别作 x 轴的垂线，分别交 $y=\ln x$ 的图象于 A'，B'，C' 三点．因为 $y=\ln x$ 是上凸函数，故有 $BB'>BB''$．图中的一些线段分别表示为：

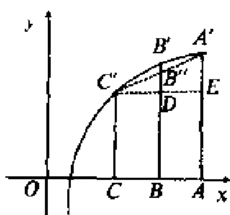


图 6－31

$$AA'=\ln a,\qquad BB'=\ln b,\qquad CC'=\ln c,$$
$$BC=b-c,\qquad AB=a-b,\qquad AC=a-c .$$

因为

$$BB'>BB''=BD+DB''=CC'+\frac{BC\cdot A'E}{AC}$$
$$=CC'+\frac{BC\cdot AA'-BC\cdot AE}{AC}.$$

于是得

$$AC \cdot BB' > AC \cdot CC' + BC \cdot AA' - BC \cdot AE$$
$$= CC'(AC - BC) + BC \cdot AA'$$
$$= AB \cdot CC' + BC \cdot AA'.$$

这个不等式意味着：

$$(a-c)\ln b > (a-b)\ln c + (b-c)\ln a,$$

或 $$b^{a-c} > c^{a-b} \cdot a^{b-c}.$$

两边同乘以 $b^c \cdot c^b \cdot a^c$，得

$$c^b \cdot b^a \cdot a^c > b^c \cdot c^a \cdot a^b.$$

三、图论知识的简单运用

数学研究的对象，除了数量关系之外，还有许多别的关系．它们的状态有时可用“图”来描述．这里所说的图是一些由点与线组成的图形，即图论中所说的图．但图的某些基本概念，并不需要太多的专门知识，却可以用来解一些数学题，形成一种独特的解题思想．由于“图”具有一定的直观性，也能起到一些形数结合的作用，所以我们把它放在数形结合思想这一节里讨论．

利用图论知识解题，也是集转化与构造于一体的方法．先根据提出的问题构造出一个数学模型——图，再把要研究的问题转化为图的问题，从对图的性质的研究导出问题的结论．

1．利用图的基本概念解题

图的一些基本概念并不复杂，基本上不需要多少数学预备知识就能理解这些概念．但这些基本概念虽然简单，却是从许多离散数学对象的共同属性中抽象出来的，它反映各种数学对象之间的相互关系．因此，许多离散数学问题，都可以转化为图论模型，这些模型只涉及基本概念，却有利于发现解题方法．

例 1 有 n（$n \geqslant 3$）个人，他们之间有些人互相认识，有些人互相不认识，而且至少有一个人不与其他人都认识．问：互相都认识的人数的最大值是多少？（美国数学竞赛题）

作一个图，用 n 个点表示 n 个人，当且仅当两个人互相认识时，两点间连一条边，就得到一个图 G．本题相当于问在这个图中最大的完全子图有多少阶．

由于至少有一人（例如 V_1）与其他人不都相识，V_1 至少与另一顶点 V_2 不相邻．因此图中最大的完全子图不能包含顶点 V_1，V_2，最多是一个 $n-2$ 阶的完全图．即互相都认识的人至多有 $n-2$ 个．

例 2 我国民间有一个流传广泛的过河趣题：某人带了一只羊，一头狼，一筐菜要过河去，由于渡船太小，人每次只能带一样东西过河，问应如何安排，才能用最少的次数渡过河去？

在过河过程中，岸上会出现各种状态，有些状态是容许的，如（人，羊），有些状态是不容许的，如（羊，菜）．容许的状态共有10种，如图 6－32 所示．把它们用点表示作为图的顶点．若从一个状态通过一次渡河可以变为另一个状态，则在这两个状态之间连一条边，否则不连．例如当原岸是（人，狼，羊，菜）时，人带羊过河去后，原岸留下状态（狼，菜），故这两点之间连一条边．但不能再通过一次渡河使原岸由（狼，菜）变为（人，狼，羊），故这两点间不连线．这样就得到一个图．

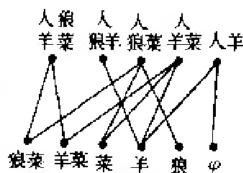


图 6－32

本题的问题是：是否存在从（人，狼，羊，菜）到 φ 的一条道路？这条道路最短有多长？由图 6－32 知，图中有两条

道路，它们的长度都是7．即存在两种过河的方法，每种都要摆渡7次．

2．图的染色

图的染色是在构造了模型的基础上再运用分类的思想．在两者的配合之下，形成一种独特的数学思想．

例1 在图上任取$n(>2)$个点，把每个点用线段与其余各点相连接，能否一笔画出所有这些线段，使第一条线段的终点与第二条线段的起点相重合，第二条线段的终点与第三条线段的起点相重合，…，最后一条线段的终点与最初那条线段的起点相重合？

这是典型的欧拉图问题．以n个点为图的顶点，从每一个顶点可作出$n-1$条边，构成一个n阶完全图K_n．如果可以一笔画，每个顶点的度数必须为偶数，即$n-1$为偶数，n为奇数．故当n为奇数时，本题有解；当n为偶数时，本题无解．

例2 6个人参加一个集会，每两个人或者互相认识或者互不认识．证明：存在两个集合，每个集合有3人组成，在同一个集合中，成员互相认识，或者互不认识．

本例是1991年加拿大数学奥林匹克选拔试题，可用图染色来解决．

用平面上6个点表示6个人，作完全图K_6，将两人认识的边染上红色，两人不认识的一边染上蓝色．要证明图中存在两个同色的三角形．如果从一点发出的两条边同色，称它为同色角，否则称为异色角．一个同色三角形有3个同色角，一个异色三角色只有一个同色角．现在设图中有x个同色三角形，要证$x\geqslant 2$．同色角的个数为：

$$3x+(C_6^3-x)=20+2x.$$

另一方面，从顶点 A 出发的 5 条边中，如有 r 条红色，$5-r$ 条蓝色，则有 $C_r^2+C_{5-r}^2\geqslant C_2^2+C_3^2=4$，即从任一顶点出发至少有 4 个同色角，图中同色角的个数至少有 $4\times6=24$ 个．从而

$$20+2x\geqslant24,$$

即 $x\geqslant2$．